

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Scienze Biomediche: Progetto n. 1 "Biotecnologie Mediche"

Ciclo XXV

Settore Concorsuale di afferenza: 06A3

Settore Scientifico disciplinare: MED07

TITOLO TESI

**“Cell Host-Microbe Interactions: Turning Pathogen
Mechanisms Into Cell's Advantages”**

Presentata da: Dott. Simone Avanzi

Coordinatore Dottorato

Relatore

Prof. Lucio Cocco

Prof.ssa Marialuisa Zerbini

Esame finale anno 2013

INDEX	2
Introduction	3
Abstract	5
Chapter I - Chlamydiaceae Moves to the Nucleus: Whole Proteome Analysis Scanning for Nuclear Localization Signals	6
Chapter II - Susceptibility of Human Placenta Derived Mesenchimal Stromal/Stem Cells to Human Herpesviruses Infection	23
Chapter III - Turning the HIV integrase into a dis-integrase: tailoring a viral enzyme to promote the disruption of its natural target	50

INTRODUCTION

Describing host-pathogen interactions in few pages is far to be exhaustive and it is not my intent. With this brief introductory part I would like to take into account several well-known host-pathogen interactions. My hope is that this would help the reader to understand how sophisticated are the strategies adopted from pathogens to take advantage of host machinery.

Host pathogen interactions.

In evolutionary terms pathogenesis is the true “world’s oldest profession”, since host and pathogens have become to “negotiate” from the first event of phagocytosis. Starting from that point host cells began to evolve, developing new metabolisms, organizing in more complex structures (tissue, organ, organism), and erecting defenses from pathogens attack. The more specific and selective became host defenses, the more pathogens had to evolve specific countermeasures. This challenge is still working leading to an unceasing co-evolution of the two counterparts. Beside evolutionary aspects, this linked evolution have created strict connections between host and pathogens proteins: it is possible to infer the function of some pathogens’ protein in the case its host interacting protein has been identified, or understanding the activity of some pathogens’ protein could lead to uncover unknown cell activity. In fact many research advances in eukaryotic cell molecular biology have been done thanks to experimental research on cell parasites: for example Kalderon and coworkers were studying SV40 large T-antigen protein, usually located inside nuclear compartment of expressing cells, while they uncover a T-antigen variant diffusely distributed inside expressing cell. The sequence analysis of this protein variant led to the identification of a short sequence capable of conferring nuclear localization to otherwise cytoplasmic proteins: they identified the first Nuclear Localization Sequence (NLS), an eukaryotic signal adopted by cell to address to nuclear environment nuclear acting protein (Kalderon et al. ’84. *Cell*;39: 499). Similarly a few year later Malim and colleagues uncover the opposite mechanism, namely sequence mediated nuclear export, while they were working on HIV-1 Rev protein (Malim et al’ 89; *Nature* 338:254). Other cellular mechanism have been understood investigating mostly viral activity, leading to a deeper knowledge of the cell at a molecular level (Knipe and Holley, *Fields virology*, 5th edition). Indeed one of the cellular process firmly necessary for viral replication is protein translation: no known virus codes a full set of enzyme required for protein translation. Eukaryotic protein translation is an highly controlled process, which could be represented essentially as a three step process namely initiation, elongation and termination. Each process involves a different set of proteins, which availability is regulated by intracellular or extracellular signals. The bottleneck process is initiation process in which the 3’ and 5’ untranslated region of cellular mRNA are specifically recognized by initiation factors, namely eIF4F complex which recognize 5’ CAP structures and PABP a polyA binding protein whose accessibility is regulated by multiple phosphorylation pathway. Those proteins are also necessary for the docking to ribosome of the mRNA. Initiation step is obviously targeted by virus with the purpose of overwhelm eukaryotic specific mechanisms: influenza virus for in-

stance steals host 5' CAP sequences together with a proximal twenty nucleotides host mRNA fragment, that is used to prime viral RNA synthesis. Moreover viral NS1 protein interacts with eIF4F complex to recruit it to viral mRNA (Plotch et al. '81 Cell 23: 847 , Dias et al. ' Nature 458: 914 , Burgui et al. ' 07 j Virol. 81: 12427) resulting in a strong translation inhibition for most of the cellular mRNAs. Other viruses adopt a mimic strategy by encoding mRNA provided with 5' CAP and polyA sequences. This is the case of herpesviruses which protein translation rely on CAP-dependent translation; in fact HSV1 stimulates eIF4F availability, by a activating degradation of its inhibitory subunit. Furthermore HSV1 encodes the multifunctional ICP6 proteins, which strengthen mRNA docking to initiation complex (Walsh et al.' 06 Genes Dev. 20: 461 , Walsh et al. ' 04 , Genes Dev 18: 660, Walsh et al. ' 08 , j Virol. 28:2648). The above described examples are exquisite representation of the intricate and intriguing field of host-pathogen interaction.

The growing knowledge of pathogens-cell interaction at molecular level, as wells as the growing technologies capability, supporting research field, are laying the foundations for the development of innovative therapeutic approaches: pathogens proteins, or pathogens themselves are engineering in order to create molecular tool to counteract different diseases. A fascinating approach has been illustrated very recently in our city of Bologna by Campadelli-Fiume virology research group: they have engineered HSV-1 virus in order to redirect its specific receptor recognition to another receptor, namely epidermal growth factor receptor type 2 (HER 2), highly expressed in high-grade glioma malignancy. Their modified virus are capable to tackle tumor growth, tumor proliferation and metastases spread in mice models (Nanni et al. 13, Menotti et al. ' 12).

Abstract

Host-Pathogen Interaction is a very vast field of biological sciences, indeed every year many unknown pathogens are uncovered leading to an exponential growth of this field. The present work lies between its boundaries, touching different aspects of host-pathogen interaction: We have evaluated the permissiveness of Mesenchymal Stem cell (FM-MSC from now on) to all known human affecting herpesvirus. Our study demonstrates that FM-MSC are fully permissive to HSV1, HSV2, HCMV and VZV. On the other hand HHV6, HHV7, EBV and HHV8 are susceptible, but failed to activate a lytic infection program. FM-MSC are pluripotent stem cells and have been studied intensely in the last decade. FM-MSC are employed in some clinical applications. For this reason it is important to know the degree of susceptibility to transmissible pathogens. Our attention has then moved to bacterial pathogens: we have performed a proteome-wide *in silico* analysis of *Chlamydiaceae* family, searching for putative Nuclear localization Signal (NLS). *Chlamydiaceae* are a family of obligate intracellular parasites. It's reasonable to think that its members could deliver to nucleus effector proteins via NLS sequences: if that were the case the identification of NLS carrying proteins could open the way to therapeutic approaches. Our results strengthen this hypothesis: we have identified 72 proteins bearing NLS, and verified their functionality with *in vivo* assays. Finally we have conceived a molecular scissor, creating a fusion protein between HIV-1 IN protein and FokI catalytic domain (a deoxyendonuclease domain). Our aim is to obtain chimeric enzyme (trojIN) which selectively identifies IN naturally occurring targets (HIV LTR sites) and cleaves subsequently LTR carrying DNA (for example integrated HIV1 DNA). Our preliminary results are promising since we have identified trojIN mutated version capable to selectively recognize LTR carrying DNA in *in vitro* experiments.

Chapter I

Chlamydiaceae Moves to the Nucleus:
Whole Proteome Analysis Scanning for
Nuclear Localization Signals

Chlamydiaceae Moves to the Nucleus: Whole Proteome Analysis Scanning for Nuclear Localization Signals

Abstract.

Chlamydiaceae family are composed by obligate intracellular bacteria. Like other intracellular parasite they exploit cell machinery to their advantage. To date lot of Chlamydiaceae effector proteins have been identified acting inside host cytoplasm environment, but only a single recent work reported a nuclear acting protein. We believed that other unidentified Chlamydiaceae proteins could be redirected to nuclear compartment during infection. To confirm our hypothesis we performed an *in silico* analysis on each Chlamydiaceae proteome in order to identify putative Nuclear Localization Signals (NLS): 72 proteins representing the 1% of the total amount have one or more putative NLS embedded in their sequence. To test the predictively of our system we have analyzed the ability of two identified NLS to confer nuclear localization to the normally ubiquitous protein, GFP.

INTRODUCTION

The family *Chlamydiaceae* comprises a diversified group of obligate intracellular Gram-negative bacteria infecting a wide range of different cell types in their eukaryotic hosts, including those of the eye, lung and genital tract epithelium, causing a variety of acute and chronic diseases (1,2,3,4). *Chlamydiaceae* are characterized by a biphasic cycle of development with spore-like, infectious Elementary Bodies (EB) and intracellular dividing, metabolically active Reticulate Bodies (RB), that inhabit a specialized parasitophorous (5) non-fusogenic inclusion separated from host cell cytoplasm by a double layered inclusion membrane (6). *Chlamydiaceae* have evolved adaptation strategies that fit them for thriving in a particularly hostile environment, the inside other cells: during infection, a temporally regulated gene expression program hits the host cell with successive waves of effector proteins that modulate host cellular functions in coordination with the chlamydial infectious cycle and/or in response to cues from the host cell. Some of these effectors appear to significantly impact the cell cycle of infected cells, with evidence for cleavage of the mitotic cyclin B1 (7), delays in cytokinesis (8) and supernumerary centrosomes (9). It would be reasonable to expect that part of these host-interactive proteins may exert their function inside the cell nucleus, an easy prediction that is supported by a recent report showing that NUE, a *Chlamydia trachomatis* protein, migrates into the nucleus and methylates histones (10). Active nuclear import of chlamydial proteins has yet to be documented, although it has already been described previously for other bacterial proteins, leaving the open possibility to uncover Chlamydial NLS: examples of prokaryotic proteins imported in cell nuclei are the *Shigella flexneri* effector OspF (11) and IpaH9.8 (12), *Yersinia* spp. YopM (13), *Xantomonas* spp. AvrBs3-like proteins (14), VirD2 protein in *Agrobacterium tumefaciens* (15), the IgA1 protease in *Neisseria gonorrhoeae* (16) and the Ribosomal proteins in *Thermus thermophilus* (17).

PY-NLS	GFP	1385 VGPVRPTFYALHFNPY 1400
AA-NLS	GFP	1385 VGPVRPTFYALHFNDY 1400
DY-NLS	GFP	1385 VGPVRPTFYALHFNAA 1400
Bip-NLS	GFP	790 KKRECAGGAIFAKRVR 805
K+ NLS	GFP	391 PTTKRGRSGGEDARADALKKPK 413

FIG. 1

Schematic representation of the constructs used in this study. Analyzed putative NLS sequences are listed in white boxes. The NLS coordinates in protein primary structure are

Nuclear import is a well regulated active pathways which is mediated by members of the karyopherin family, a large, evolutionarily conserved family of transport factors. Karyopherins interact directly with their cargoes, although some also use adapter proteins. Karyopherins that mediate import bind to their cargoes in the cytoplasm via recognition of the nuclear localization signal (NLS), a short sequence embedded in the primary structure of cargo proteins. The karyopherin: cargo complex translocates through the NPC via interactions with NPC proteins (18). NLS signals were first identified and characterized on the SV40 Large T antigen and Nucleoplasmin proteins, as one or two clusters of basic aminoacid residues separated by a linker, respectively (19) (see Tab. 1). The nuclear import of many proteins is mediated by these so called classic NLS. However evidence has accumulated for the existence of additional, non-classic import signals unrelated to the basic NLS (18). In some cases, the NLS contains several basic amino acids as has been determined for core histones and ribosomal proteins. In other cases, the NLS domain is relatively large, as in the case of M9 NLS, first identified in hnRNP A1 by Siomi and Dreyfuss (20). This non classic NLS consists of about 38 amino acids, is glycine rich, deficient in basic amino acids and its nuclear transport is directly mediated by Karyopherin- β 2. Recently Lee and coworkers have determined the M9NLS sequence necessary and sufficient for proper Karyopherin- β 2 recognition and subsequent nuclear import. Furthermore they found that similar non classic sequence is found in other known Karyopherin- β 2 binding proteins. Crystallographic data of M9NLS peptide bound to Karyopherin- β 2 together with mutagenesis analysis supported them in defining consensus sequences for M9 like NLS, reported in Table 1:

NLS prototype	SubType	Consensus
M9NLS/PY-NLS	Basic	[KR]-X(0,2)-[KR]-[KR]-X(3,10)-[RKH]-X(1,5)- P-Y
	Hydrophobic	[AILMFWYV]-[GAS]-[AILMFWYVPKQR]-[AILMFWYVPKQR]-X(7,12)-[RKH]-X(2,5)- P-Y
Canonical NLS	Monopartite	PX(1,3)[KR][KRH](0,1)[KR][KR][KR]
	Bipartite	K-[KR]-X(10,12)-K[KR]X(0,1)[KR]

Tab. 1 Consensus sequences for canonical and non canonical Nuclear Localization Signals (NLS) under investigation in this work

M9NLS is a long and structurally disordered substrate, with an overall positive charge (or neutral at least). Its central domain could be hydrophobic or basic, defining the two subtypes of this NLS. Finally it has C-terminal basic residue followed by an invariant PY sequence, after which this kind of NLS is named. PY-NLS consensus could be predictive for identifying unknown Karyopherin- β 2 binding targets,(21). A recent work has observed a putative Bipartite NLS signal in *Chlamydia trachomatis* pmpD protein (22), supporting the hypothesis that some *Chlamydiae* protein actively transfer into their host nuclear compartment. We report here for the first time an *in silico* whole proteome analysis on *Chlamydia* and *Chlamydiophila* species in order to identify putative NLS. In addition we have probed experimentally verified the effective nuclear import of GFP protein fused either to the previously reported Bipartite-NLS or to a PY-NLS, that we have identified in the present work, both occurring on the *Chlamydia trachomatis* pmpD protein.

Material and Methods

Sequences Genesis. All mammalian constructs expressing GFP fusion proteins were generated using the Gateway system (Invitrogen, Carlsbad, CA). Oligonucleotides including the attB1 and attB2 recombination sites were used as both primers and template to generate the sequences of interest by fill-in reaction with GoTaq Taq polymerase (Promega, Madison, WI) as follows: primers pmpDattB1 (5'-GGGGACAAGTTTGTACAAAAAAGCAGCAGGCTTCGTTGGTCCTGTGAGACCTACTTTTATGCGCATTTCAT) and pmpD-PYattB2 (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAATAAGGATTGAAATGCAAAGCA TAAAAAGTAGG) were used to obtain PY fragment, oligonucleotides pmpDattB1 and pmpD-DYattB2 (5'-GGGG ACCACTTTGTACAAGAAAGCTGGGTCTTAATAGTCATTGAAATGCA AAGCATAAAAAGTAGG) were used to generate DY fragment, primers pmpDattB1 (5'-GGGGACAAGTTTGTACAAAAAAGCAGCAGGCTTCGTTGGTCCTGTGAGACCTACT

TTTTATGCGCATTTC AAT) and pmpD-AAattB2 (5'-GGGGACCACTTTGTACAA GAAAGCTGGGTCTTAAGCGGCATTGAAATGCAAAGCATAAAAAGTAGG) were used to generate AA fragment. Primers pmpD-BiPattB1 (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAAAAGAAGAGAGTGTGCTGG) and pmpD-BiPattB2 (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCTT AACGAACCCGTTTTGCAAAAAT) were used to generate BiP sequence using purified Chlamydia Trachomatis serovar L2 genome as PCR template. Agarose-purified PCR products were introduced into the plasmid vector pDONOR207 (Invitrogen, Carlsbad, CA) via the BP recombination reaction, according to the manufacturer's recommendations, in order to create the entry clones pDNR-PY, pDNR-DY and pDNR-AA and pDNR-BiP. Constructs were then used to perform LR recombination reactions with the Gateway system compatible expression ("DEST") vector pEPI-DEST-GFP (23), according to products specifications, to express N-terminally tagged GFP fusion proteins. The integrity of all constructs was confirmed by DNA sequencing (PRIMM s.r.l., Milan).

Cell lines. African Green Monkey COS-7 cells were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 5% (v/v) fetal bovine serum, 50 units/mL penicillin, 50 units/mL streptomycin, and 2 mM L-glutamine. For live imaging experiments, cells were trypsinized and 2×10^5 cells were seeded onto coverslips in 24-well plates 1 day before transfection, which was performed using TransIT-LT1 (Mirus, Madison, WI) reagent according to the recommendations of the manufacturer.

ATP/GTP depletion assay. ATP/GTP pool have been depleted as reported previously by us (24): briefly 24 h after transfection COS7 expressing GFP fusion protein were treated for 30 min at 37 °C in DMEM containing no glucose, 5% fetal bovine serum, and supplemented with 10 mmol/L sodium azide and 6 mmol/L 2-deoxy-D-glucose (Sigma). After this treatment, the ATP intracellular levels were restored by replacing the energy depletion medium with DMEM, as used in normal growing conditions. Each experiment has been repeated three times.

Live imaging. Live COS-7 cells, grown on coverslip and expressing GFP fusion proteins, were imaged using a Nikon Eclipse TE2000-U inverted microscope equipped with a Nikon DXN1200 digital camera and a Nikon Plan Fluor 40 \times objective (Nikon Corp., Tokyo, Japan). Digital images were obtained using the NIS elements AR 2.30 software (Nikon Corp., Tokyo, Japan), using the following settings: gain, 50; offset, 0; acquisition times 500-2000 ms. Semiquantitative analysis of the levels of nuclear accumulation relative to each GFP fusion protein was performed using NIS elements AR 2.30 software to perform single cell measurements of the nuclear (F_n) and cytoplasmic (F_c) fluorescence, subsequent to the subtraction of fluorescence due to autofluorescence/ background, to determine the nuclear to cytoplasmic fluorescence ratio (F_n/F_c). Localization of each cell was classified on the basis of the F_n/F_c ratio as exclusively nuclear (N; $F_n/F_c > 10$), mainly nuclear (N > C; $10 > F_n/F_c > 2$), or diffuse (D; $2 > F_n/F_c > 1$). At least 100 GFP expressing cells were analyzed for each fusion protein. Each experiment has been repeated three times.

***In silico* analysis.** Bioinformatic analysis was performed essentially as described by others (21): briefly proteome sequences of *C. trachomatis* serovar D, *C. pneumoniae*, *C. psittaci*, *C. muridarum*, *C. caviae*, *C. abortus*, *C. felis*, *C. suis* and *C. pecorum*, were obtained from UniProtKB database (25). Proteins carrying putative NLS were identified by subjecting each proteome to ScanProsite software, searching NLS motifs listed in tab. 1. All resulting entries were filtered for structural disorder using DisEMBL tool (26) and for an overall positive charge surrounding NLS sequence. PY-NLS overall positive charge were determined analyzing the sequence encompassing 40 aminoacid N-terminal to PY domain and 10 aminoacid C-terminal to it. Only NLS bearing proteins sharing some degree of homology with ortholog/paralog or proteins belonging to autotransporter's protein family were selected. Ribosomal proteins, protein elongation factors, signal transducers and ion carriers were removed from NLS-bearing protein list.

Results.

NLS prediction in *Chlamydiaceae*. We have performed an *in silico* analysis on the proteome of nine *Chlamydiaceae* species, scanning for putative NLS sequences matching the reported consensus for classical and PY NLS (Tab.1) (19): we have identified 336 proteins bearing potential NLS, representing the 5% of the total analyzed proteins. Subsequently we restricted the number of NLS candidates by analyzing molecular disorder and the overall charge of the sequence surrounding each NLS; furthermore we have excluded known nucleic acid interacting proteins (see material and methods section): the resulting 72 proteins with putative NLS, representing about 1% of all analyzed proteins are listed in table 2.

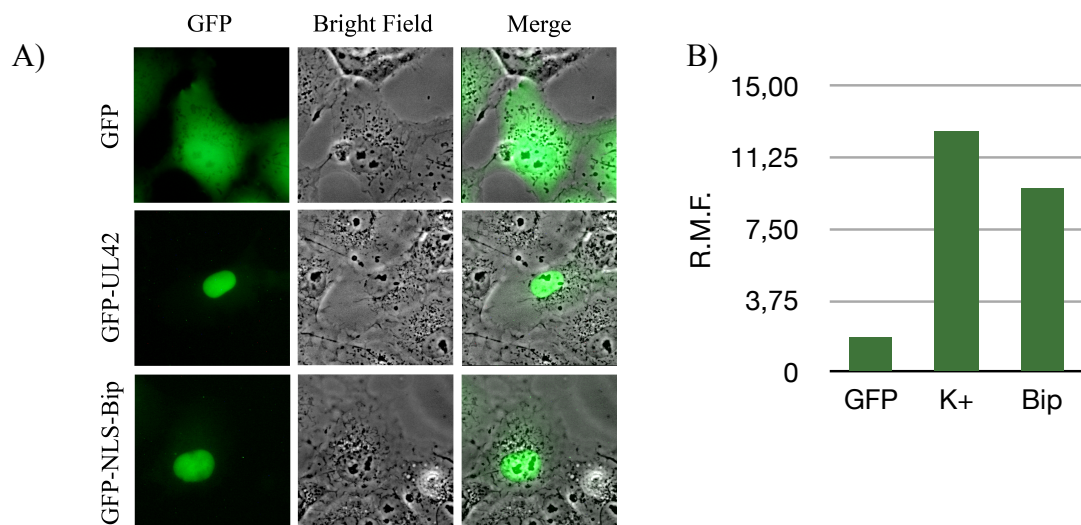


Fig.2 Bip NLS Translocate GFP into Nucleus.

- A) Live imaging analysis of COS7 cells transfected with the indicated GFP fusion protein. Pictures are representative of typical subcellular localization.
 B) Semi quantitative analysis of relative mean fluorescence (R.M.F.) of analyzed GFP fusion proteins.

Identified NLS are functional.

We have evaluated the predictability of our selection system analyzing the actual nuclear import of two identified NLS, N-terminally fused to GFP protein (see fig.1 for a schematic construct representation). Interestingly, the selected NLS are both present on the same protein, namely *Chlamydia trachomatis* pmpD protein, and respond either to classic or PY reference sequences. Our choice fell on pmpD for several reasons: pmpD belongs to the type V or autotransporter secretion pathway, the most widespread secretion mechanism employed by gram-negative bacteria to deliver virulence factors (27). Furthermore it has been suggested that late in infection, pmpD could be processed and released into the cell cytosol (28), becoming a perfect candidate for nuclear import. The canonical NLS is a putative bipartite NLS (Bip from now on), previously observed by Swanson and coworkers (22), carried by a 30kDa soluble peptide produced upon pmpD auto-processing. As clearly observable in fig. 2, Bip sequence is sufficient to confer nuclear localization to its fusion partner, the otherwise ubiquitous GFP protein. Furthermore nuclear localization of GFP-Bip fusion protein is inhibited after GTP pool depletion (fig. 3), suggesting that observed nuclear import is due to an active process (24) rather than a passive mechanism.

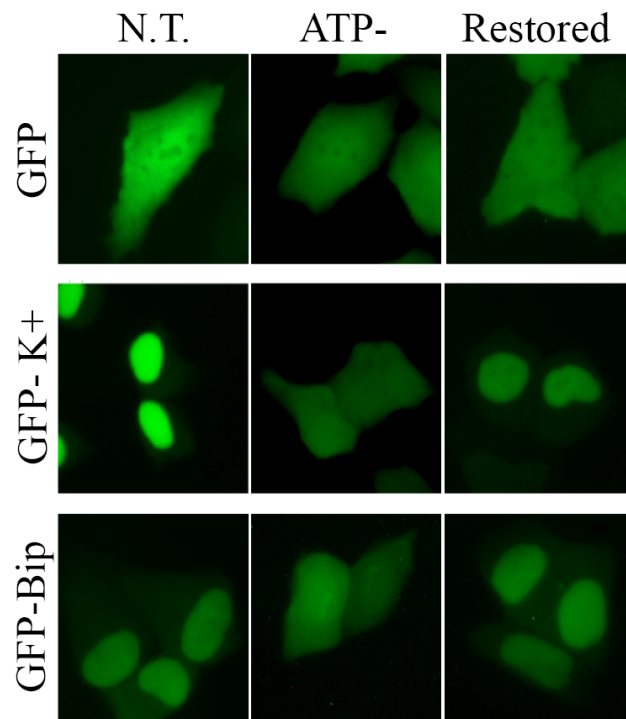
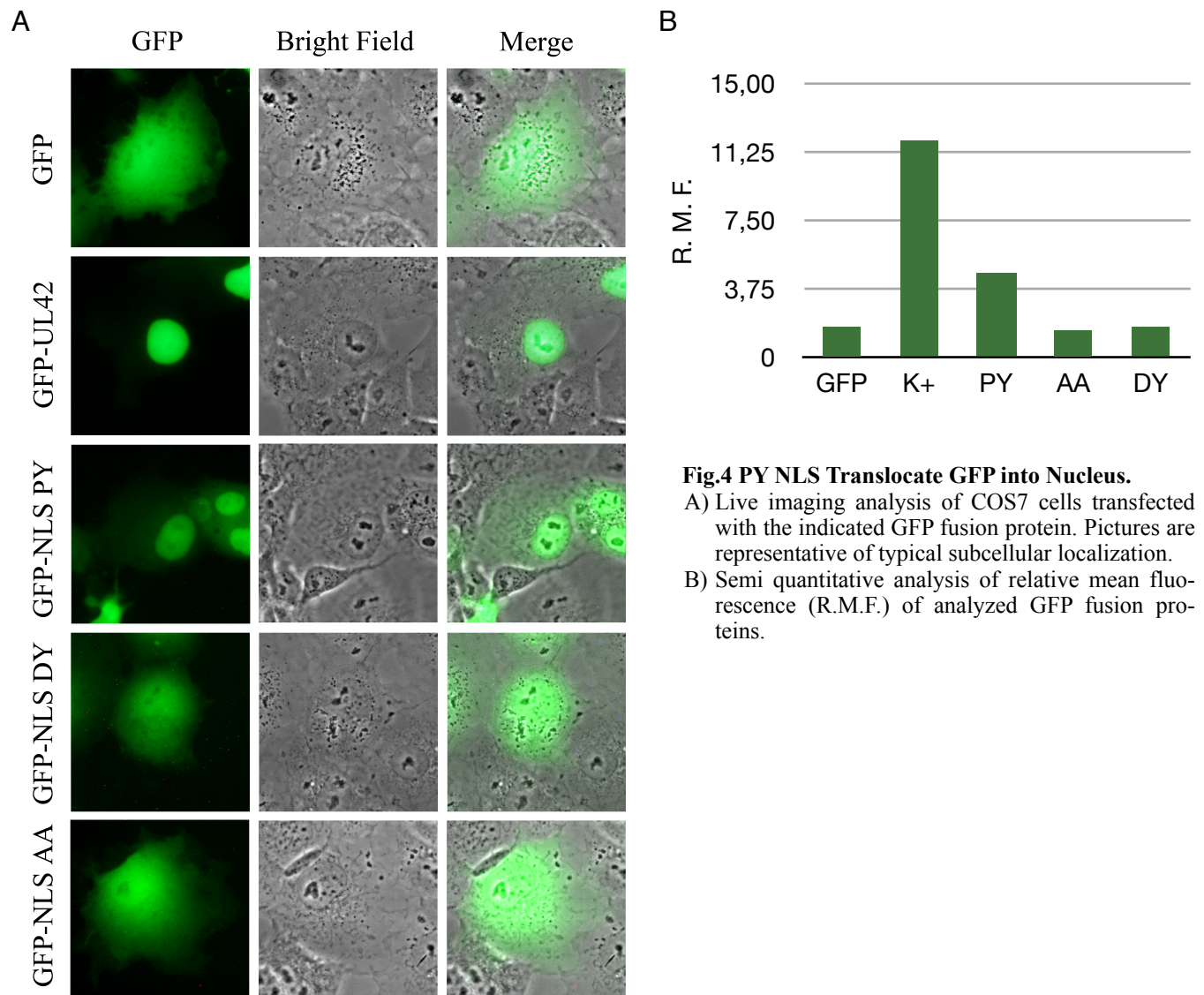


FIG. 3 Triphosphate depletion assay. Live imaging analysis of COS7 cells expressing GFP, GFP-K+ and GFP-Bip, during ATP/GTP depletion assay (ATP-) and the subsequent restoring of physiological conditions (Restore). Control untreated samples was also observed (N.T.). For each sample pictures are representative of their typical subcellular localization.

The selected PY NLS sequence is a hydrophobic PY NLS and is embedded in the pmpD C-terminal translocation domain (TD) fragment. Little is known about the pmpD TD processing fate during infection. Curiously, while the Bip NLS is not conserved in other members of the *Chlamydiaceae*, a PY NLS was also identified in two pmp paralogues of *C. pneumoniae* suggesting an unknown conserved function for this sequence. Fig.4 shows that PY-GFP fusion protein localizes in the nucleus of expressing cells. In parallel we analyzed subcellular localization of GFP-DY and GFP-AA fusion proteins, in which respectively P or PY residues are mutated in order to abolish Karyopherin- β 2 recognition sequence. Both fusion proteins show a diffuse subcellular localization, comparable with that corresponding to GFP alone. We could infer that the identified PY is a *bona fide* functional NLS, necessary to GFP nuclear import.



Discussion.

Members of the *Chlamydiaceae* family are obligate intracellular pathogens, and have evolved several strategies and mechanisms to exploit cellular machinery: in fact several chlamydial proteins are inserted into the inclusion membrane or secreted into the host cytoplasm. Those proteins are capable of interacting with members of the cellular machinery: for example they modify host cell cycle phase (7), they prevent host cell apoptosis (29), or inhibit NF-Kb signaling pathway, stabilizing Ikb protein (30). Chlamydial effector proteins working within the cell cytoplasm have been deeply documented while only a recent single report demonstrated a nuclear activity for *C.trachomatis* NUE protein which interacts with cell histone proteins: this is the first documented *Chlamydiaceae* nuclear protein (10). Although the Authors have not “formally” characterized yet the NLS required for NUE nuclear import, they demonstrated that a NUE protein deprived of two separated basic residue stretches failed to accumulate in the cell nucleus, suggesting that inside removed sequences could reside a yet unidentified functional NLS. Although a vastly unexplored field, bacterially encoded proteins that localize into cell nuclei represent an innovative field to better understand host-pathogens interplay. Moving in that direction we have performed an *in silico* analysis of all known *Chlamydiaceae* proteomes, scanning for conserved putative NLS. We have identified 72 proteins with one or more putative NLS representing the 1% of the 7066 analyzed. Intriguingly about 40% of identified NLS carrying proteins belong to the autotransporter protein family, whose members typically translocate in the pathogen outer membrane where they often release an active protein inside the cell cytoplasm by self processing of the full length protein. NLS occurring on autotransporter suggest a fascinating hypothesis: chlamydiae could have implemented an autotransporter system including effector proteins endowed with NLS, in order to allow their translocation in the nucleus of host cells, where they may interfere with the regulation of cellular functions, for example via epigenetic reprogramming, as the methylation of histones by NUE would suggest.

Furthermore we have verified the ability of two, among identified NLS, to confer nuclear localization to GFP protein, which normally appears diffuse ubiquitously in expressing cells. The selected NLS occur in two different domains of *C.trachomatis* pmpD, an autotransporter protein highly conserved among *trachomatis* serovars, which suggests the idea that pmpD exerts an essential although as yet unknown function. One NLS is a classical Bipartite NLS (19), inserted in the Passenger Domain, the secreted effector of pmpD protein. Our results demonstrated that it is capable to confer nuclear destination to GFP protein; furthermore we have demonstrated that GFP-Bip nuclear import is an active transport, which suggests that nuclear import is mediated by the Bip NLS, as reported before (24). The other selected NLS is a PY NLS which is embedded into the pmpD Translocation Domain. While the former domain is known to be processed and secreted into the cytoplasm during chlamydia infection, the latter theoretically would remain membrane associated (28,22). In contrast we have found that GFP-PY fusion protein accumulates in the nuclei of expressing cells. Furthermore mutant fusion protein with abolished PY sequence failed to be actively imported beyond the nuclear envelope, showing a diffuse subcellular localization. Taken together these results suggest that the identified NLSs are functional. In order to confirm actual NLS functionality, we need to demonstrate effective physical interaction

between GFP-NLS fusion protein and related cellular transporter (karyopherins). Furthermore it will be also necessary to identify during infection an NLS associated peptides inside the nuclear environment.

In this study we have identify 72 Chlamydiaceae proteins bearing one or more NLS signals. This could be the starting point for the discovery of additional unknown nuclear activities of Chlamydiaceae's proteins. In turn, uncovering new nuclear localizable effector functions would shed light on the complex interaction between host cell and chlamydiaceae.

Protein	Chlamydia/ Chlamydiophila species (IDprot)	NLS http://prosite.expasy.org/scanprosite/	Function/ Membrane/ Secretion http://www.uniprot.org/
		K-[KR]-X(10,12)- K[KR]X(0,1)[KR]	
10 kDa chaperonin	caviae (P15598)	KKqdraevlvtgKRdK 40 – 57	chaperon, protein folding
	psittaci (F0T4F9)	KKqdraqvltgKRdK 40 – 57	
	abortus (Q6YNR5)	KKqdraevlvtgKRdK 40 - 57	
Serine/threonine-protein kinase pknD	caviae (Q822K5) *	KRiredlsdnell.KK.R 33 – 48	protein kinase
	abortus (Q5L5J7)		
	psittaci (F0T4J3)	KRirkdlsdnell.KK.R 33 – 48	
	pneumoniae (Q9Z986) *	KKiredlaenpll.KR.R 33 – 48	
Polymorphic outer membrane protein	trachomatis serD (D7DET2) *	KRrecaggaifa..KRvR 790 - 805	autotransporter
Polymorphic membrane protein D	suis (B8YPF0) *	KRrecaggaifa..KRvR 783 - 798	autotransporter
UPF0235 protein CCA_00247	caviae (Q824A6)	KRdvtlissetr.KK.R 58 – 73	
UPF0235 protein CAB243	abortus (Q5L6M2)	KRditlpgdtsr.KK.R 58 - 73	
UPF0235 protein G5S_0576	pecorum (F4DL14)	KRdvtlvsgessr.KK.K 69 – 84	
FtsK	trachomatis serD (D7DEL2)	KRnvhrpeslleelKK.K 334 - 350	Integral to membrane, cell division.
Type III secretion chaperone SycD	caviae (Q82118) *	KKtrsrilaalaaq.KKaK 18 – 34 KKkkataaaksttdKK.K 206 – 222	response to stress
LcrH	pneumoniae (Q9Z795)	KKtrsrilaalaaq.KKaK 21 – 37	
Type III secretion chaperone SycD	Psittaci (F0T5A6)	KKtrsrilaalaaq.KKaK 18 – 34	
Putative uncharacterized protein lcrH_1	trachomatis ser D (D7DE48)	KKsrsrlaiaaqa.KKaK 18 – 34	
Protein translocase subunit, putative	caviae (Q823A4)	KKinrndpcpgsnKKyK 3 – 20	
Protein translocase subunit	pecorum (F4DIZ8)		
Protein translocase subunit	pneumoniae (Q9Z8S8)	KKinrndlpcpgsnKKyK 3 – 20	
Protein translocase subunit	psittaci (F0T453)	KKvnrndpcpgsnKKyK 3 – 20	
Putative uncharacterized protein	abortus (Q5L5X1)		
Preprotein translocase subunit-like protein	trachomatis (D7DCW3)	KKpnrndpcpgsgKKyK 3 – 20	
Hypothetical membrane associated protein	trachomatis (D7DEK7) *	KKksscyllkffsrKKpK 57 – 74	
Putative uncharacterized protein	caviae (Q824I3)	KRqlyvnrglggf.KRIR 292 – 308	hydrolase activity
Phosphohydrolase	psittaci (F0T357)		
Putative exported protein	abortus (Q5L6V1)	KKqlyvnrglggf.KRIR 292 – 308	

Putative uncharacterized protein	trachomatis (D7DEE3)	KKdrvdraekvvkeKR.R 13 – 29	
	caviae (Q824V3)		
	pneumoniae (Q9Z7J9)		
	psittaci (F0T5K1)		
	pecorum (F4DKF7)		
Putative uncharacterized protein	caviae (Q821R6) *	KKtkelsaeaeII.KKIR 4 – 20 KKIrdqalvleeqkKR.K 17 – 33	
	psittaci (F0T526) *	KKtkelsaeaeII.KKIR 4 – 20 KKIrdqalileeq.KK.R 17 – 32	
	abortus (Q5L516) *	KKydfIrfmgmt.KKIK 3 – 19 KKikelsaeaeII.KKIR 16 – 32 KKIrdqalaleeq.KK.R 29 - 44	

		PX(1,3)[KR][KRH](0,1)[KR][KR][KR]	
Serine/threonine-protein kinase pknD	pneumoniae (Q9Z986) *	Py.RRKK 229-234	protein kinase
	trachomatis ser D (D7DDC2)	Py.RRKK 235-240	
	muridarum (Q9PK92)		
	caviae (Q822K5) *	Py.RKKK 230 – 235	
General Secretion Protein F	pneumoniae (Q9Z789)	PkeRRKR 11 - 17	Type II protein secretion system
Uncharacterized protein TC_0273	muridarum (Q9PL35)	PspRRKK 355 - 361	Integral to membrane
Preprotein translocase, YajC subunit	pecorum (F4DJZ1) *	PeqKRRK 51 - 57	Integral to membrane
UPF0092 membrane protein CPn_0884	caviae (Q821Q4) *	PeqKRRK 52 - 58	
	psittaci (F0T538) *		
	pneumoniae (Q9Z722)		
Conserved membrane protein	abortus (Q5L504) *		
SET domain containing protein	pecorum (F4DJZ6)	PlywKHRKK 201 - 209	hystone metilation
	trachomatis ser D (D7DEL0)	PlywKHRKK 202 - 210	
	pneumoniae (Q9Z728)	PlywKHRKK 204 - 212	
Putative uncharacterized protein	psittaci (F0T543)	PlywKHRKK 203 - 211	
	abortus (Q5L4Z8)		
	caviae (Q821Q0)		
	muridarum (Q9PLI9)	PlywKHRKK 202 - 210	
Negative regulator of genetic competence clpC/mecB	trachomatis ser D (D7DDA7)	PealR.KKR 257 - 264	chaperone, nucleotide-excision repair
Endopeptidase Clp ATP-binding chain clpC	pecorum (F4DIH4)	PealR.KKR 255 - 262	
Probable ATP-dependent Clp protease ATP-binding subunit	muridarum (Q9PKA8)	PealR.KKR 274 - 281	
Negative regulator of genetic competence clpC/mecB	pneumoniae (Q9Z8A6)	PdalR.KKR 255 - 262	
	abortus (Q5L6D0)	PdtlR.KKR 254 - 261	
ATP-dependent Clp protease ATP-binding subunit	caviae (Q823Q2)		
	psittaci (F0T3N8)		

Putative uncharacterized protein	trachomatis ser D (D7DCN8)	PrgsK.KRR 106 - 113	
	abortus (Q5L656)	PrgsK.KRR 86 - 93	
	muridarum (Q9PKX4)	PrgsK.KRK 83 - 90	
Putative uncharacterized protein CPj0350	pecorum (F4DIQ7)	PrgsK.KRR 83 - 90	
	caviae (Q823H4)		
	psittaci (F0T3W8)		
	pneumoniae (Q9Z8J3)		

		[KR]-X _(0,2) -[KR]-[KR]-X _(3,10) -[RKH]-X _(1,5) -P-Y	
Hemolysin (Putative uncharacterized protein CPj0395)	pneumoniae (Q9Z8E8)	KdhKKaryinfl...Rr....PY 43 - 59	FAD-binding
	psittaci (F0T3T4)		
	caviae (Q823K9)		
	pecorum (F4DIM1)		
CBS domain protein, putative			
CBS domain protein			
CBS domain protein	abortus (Q5L690)		
Putative membrane transport protein			
Peptide ABC transporter, periplasmic peptide-binding protein, putative	muridarum (Q9PKJ4)	Ke.KKdrwkle..Kn....PY 201 – 216	transporter activity
Peptide ABC transporter, permease protein	muridarum (Q9PKJ3)	LAYAgklsplkvlik.Hvl...PY 207 – 226	Transmembrane, transporter activity
	psittaci (F0T4C2)	LAYAgklsplikvlik.Hil...PY 207 - 226	
Peptide ABC transporter, permease protein	pecorum (F4DJ73) *	KknKKnnkqphtqe.Kan...PY 327 - 345	Transmembrane, transporter activity
Oligopeptide Permease	pneumoniae (Q9Z8Y5) *	LAYAgklsplkvlik.Hil...PY 207 - 226	Transmembrane, transporter activity
Oligopeptide ABC transporter permease protein	pneumoniae (Q9Z7V7) *	WGQPedlcianvekRae...PY 270 - 290	Transmembrane, transporter activity
OppA	trachomatis ser D (D7DD20) *	Ke.KKdrwkle..Ks....PY 199 - 214 KkdRRwkle.....Ks....PY 201 - 214 Kd.RRwkle.....Ks....PY 202 - 214	Transporter activity
Putative ABC transporter, ATP-binding component	abortus (Q5L6G4)	ReiKKlqpqelk...KsniqrPY 288 - 307	ABC type 1 transporter
	psittaci (F0T3K3)	ReiKKlqpqelk...KsniqrPY 284 - 303	
	pneumoniae (Q9Z9F8)		
	muridarum (Q9PK46)		
Peptide ABC transporter, ATP-binding protein	pecorum (F4DIE1)		
	caviae (Q823U0)	ReiKKlqpqelk...KsniqrPY 296 - 315	
ABC Transporter Protein ATPase	muridarum (Q9PLN7)	AGRMaeyssvqdif...Hspah.PY 232 - 252	peptide transporter activity

		[AILMFWYV]-[GAS]-[AILMFWY VPKQR]-[AILMFWYVPKQR]-X(7,12)-[RKH]-X(2,5)-P-Y	
Probable outer membrane protein pmp19	pneumoniae (Q9Z813)	LSLRssfsytedsqvmKhvf..PY 758 - 779	Cell outer membrane, Autotransporter
Probable outer membrane protein pmp2	pneumoniae (Q9Z3A1)	ASLYsvvsillgeggIReill.PY 619 - 641	Cell outer membrane, Autotransporter
Polymorphic outer membrane protein pmpD	trachomatis ser D (D7DET2) *	VGPVrptfyal.....Hfn...PY 1385 – 1400	Autotransporter
Polymorphic membrane protein D	suis (B8YPF0) *	VGPVrptfyal.....Hfn...PY 1377 - 1392	
Probable outer membrane protein pmpD	muridarum (Q9PLB0)	VGPArpkfyal.....Hfn...PY 1367 - 1382	
Probable outer membrane protein pmpA	muridarum (Q9PJY3)	ASLPnnnsilv.....Kes...PY 408 - 423	Autotransporter
Outer membrane protein, putative	muridarum (Q9PKF3)	VSFAvhptlpvydvtvRvsegaPY 333 - 356	Outer membrane
Hypothetical membrane associated protein	trachomatis ser D (D7DD25)	VALlqelipt.....Hvn...PY 145 - 160	-
CBS domain containing protein	trachomatis ser D (D7DD78)	VGVPalttlvlgeivpKviai.PY 92 - 114	FAD-binding
Putative uncharacterized protein	muridarum (Q9PKD8)	VGVPalttlvlgeiipKviai.PY 92 - 114	
CBS domain-containing protein	psittaci (F0T3T5)	VGFPalttlvlceilpKaval.PY 92 - 114	
Putative membrane transport protein	abortus (Q5L689)		
CBS domain protein	pecorum (F4DIM2)	VGIPltitlifgevlpKaval.PY 91 - 113	
Predicted permease	trachomatis ser D (D7DEV9)	ISlsslqequivsyia.Kdv...PY 26 - 45	integral to membrane
Putative uncharacterized protein	muridarum (Q9PL80)		
Conserved hypothetical membrane protein	psittaci (F0T4S8)		
	abortus (Q5L5B1)		
YjgP/YjgQ family protein	caviae (Q822B8)		
YjgP/YjgQ family protein	pecorum (F4DJN7)		
Putative uncharacterized protein	pneumoniae (Q9Z6R3)	ISlsslqequivayia.Kdv...PY 26 - 45	
Arginine/agmatine antiporter	trachomatis ser D (D7DDJ6)	LSLVrpdlkagiytysRegfg.PY 67 - 89	Cell inner membrane, aminoacid transport
	muridarum (Q9PK20)	LSLVrpdlttgimysRegfg.PY 67 - 89	
	psittaci (F0T4P4)		
	caviae (Q822F2)		
	abortus (Q5L5E6)		
Chaperone protein DnaK	trachomatis ser D (D7DDL8)	IGRKfsevesei....Ktv...PY 81 - 97 78 - 94	chaperone, protein folding
	muridarum (P56836)		
	psittaci (F0T3D1)	IGRKysevesei....Ktv...PY 78 – 94	
	abortus (Q8GH79)		
	caviae (Q824B2)		
OppB	trachomatis ser D (D7DD21)	LAYAkglsfpkvliik.Hil...PY 207 - 226	integral to membrane, transporter activity
OppD	trachomatis ser D (D7DEG3)	AGRMAeyssvqeif..Hspah.PY 232 - 252	ABC transporter

Transcription-repair coupling factor	trachomatis ser D (D7DEMO)	AALKhkqkpmfhdhdiKief..PY 940 - 961	DNA repair
	psittaci (F0T531)	AALKnnsspllfnddvKief..PY 940 - 961	
	abortus (Q5L511)	AALKnnsspllfnddvKief..PY 942 - 963	
	caviae (Q821R1)	AALKnntpllfnddvKief..PY 942 - 963	
	pecorum (F4DJY3) *	AALKthsspsllqdtiKief..PY 926 - 947 LSIQqthgkneli...Kqtl..PY 1024 - 1042	
Virulence plasmid protein pGP6-D-related protein	muridarum (Q9PJF7)	FSAWllitygn.....Rqt...PY 110 - 125	
	psittaci (F0T5B3)		
PGP6-D chromosomal paralog	caviae (Q821I1)	FSAWllitygn.....Rqt...PY 110 - 125	
2-oxoisovalerate dehydrogenase alpha subunit	trachomatis ser D (D7DDG2) *	AGVVaakslipg....Kdwaf.PY 52 - 70 ASFLarttqnhsag..Rmm...PY 89 - 107	Oxidoreductase
Oxoisovalerate Dehydrogenase Alpha/Beta Fusion	pneumoniae (Q9Z9E8) *	AGVLagkslipg....Kdwsf.PY 52 - 70 ASFLarttqnhsa..Rmm...PY 89 - 107 YAYLdapirrlggl..Hapv..PY 634 - 653	
Putative oxidoreductase	abortus (Q5L6F6) *	VGVVagksllrg....Kdwsf.PY 52 - 70 ASFLgrlaphnssg..Rmm...PY 89 - 107 YAYLdapirrvagl..Hapv..PY 634 - 653	
	psittaci (F0T3L1) *	VGVVagkslipg....Kdwsf.PY 52 - 70 ASFLarlaphnssg..Rmm...PY 89 - 107 YAYLdapirrvagl..Hapv..PY 634 - 653	
2-oxoisovalerate dehydrogenase, E1 component, beta subunit	muridarum (Q9PK54) *	AGVVaakslipg....Kdwaf.PY 52 - 70 ASFLarttqnhsdg..Rmm...PY 89 - 107	
	caviae (Q823T1) *	AGVLagkslipg....Kdwsf.PY 52 - 70 ASFLarlaphnssg..Rmm...PY 89 - 107 YAYLdapirrvagl..Hapv..PY 634 - 653	
2-oxoisovalerate dehydrogenase, E1 component, alpha and beta subunit	pecorum (F4DK12) *	AGIVagksllpe....Kdwsf.PY 24 - 42 VAAFIartssnhssg..Rmm...PY 60 - 79	
Cell shape-determining protein MrdB, putative	muridarum (Q9PLK0)	LGMLcslifsgiiip.Hdtvk.PY 199 - 220	Integral to membrane, cell cycle
Cell shape-determining protein MrdB	caviae (Q821N9)	LGMLcslifsgiiis.Hekvk.PY 199 - 220	
Putative cell shape-determining protein	psittaci (F0T554)	AGMLcslifsgiiis.Hekvk.PY 178 - 199	
ABC ATPase	pneumoniae (Q9Z7M2)	AGRMvecapavqmf..Hnpsh.PY 232 - 252	peptide transporter activity
Oligopeptide ABC transporter, ATP-binding protein	psittaci (F0T5M4)	AGRLaecapvkdif..Hnpch.PY 232 - 252	
Peptide ABC transporter, ATP-binding protein	caviae (Q824T1)	AGRKaeyasakeif..Hnpch.PY 232 - 252	
Peptide ABC transporter, ATP-binding protein	psittaci (F0T4C5)	LSVLrlvnlplsvl..Rlk...PY 116 - 134	ABC transporter
RodA	trachomatis ser D (D7DEJ9)	LGMFcslifsgiiip.Hdkvk.PY 199 - 220	Integral to membrane, cell cycle
Oligopeptide Permease	pneumoniae (Q9Z8Y5) *	LAYAkglspkvkvik.Hil...PY 207 - 226	Integral to membrane, transporter activity
Oligopeptide ABC transporter permease protein	pneumoniae (Q9Z7V7) *	WGQPediciayanvekRae...PY 270 - 290	Integral to membrane, transporter activity
Type III secreted protein SctW	pneumoniae (Q9Z8L4)	MAIVssflmkgmatelKrqg..PY 233 - 254	Outer membrane, type III protein secretion system
Oligopeptide ABC transport system substrate-binding protein oppA	pneumoniae (Q9Z8Y7)	LAFPafaifkpenp..Klfsg.PY 155 - 175	Transporter activity
Rod Shape Protein	pneumoniae (Q9Z739)	IGLIgslifsgivis.Hqkvk.PY 235 - 256	Integral to membrane, cell cycle
Putative rod shape-determining protein	abortus (Q5L4Y7)	MGMLcslifsgmis.Hekvk.PY 199 - 220	
CT131 homolog-(Possible Transmembrane Protein)	pneumoniae (Q9Z8Z5)	FSLPlidsitkygkqvRls...PY 627 - 647	Integral to membrane

UvrABC system protein A	trachomatis (O84337)	ASALeistemleslft.Righ...PY 1033-1053	ABC type 1 transporter (2 domains)
Cationic Amino Acid Transporter	pneumoniae (Q9Z954)	ASLAptavlyhsntkqKel...PY 112 - 132	Membrane
Putative uncharacterized protein	muridarum (Q9PJW7)	LGRPlskgcrsdccqKkksc.PY 29 - 51	Membrane
Putative uncharacterized protein	trachomatis serD (D7DCK8)	FALLvsvsffgcq..Kli...PY 63 - 81	-
Uncharacterized protein TC_0306	muridarum (Q9PL02)	LALLvsvisffgcr..Kli...PY 64-82	transmembrane
Putative zinc metalloprotease CPn_0344/CP_0416	pneumoniae (Q9K275)	FSRQprvlasvl....Hyt...PY 281-297	transmembrane, metalloprotease
Uncharacterized protein CPn_0442	pneumoniae (Q9Z8A1)	LGLLmsvtstvlfqvHknlfPY 122-145	transmembrane
60 kDa chaperonin 2	psittaci (F0T5D4)	LGPQGscvikkd...Rvs...PY 30 - 47	chaperone, protein folding
	caviae (P59698)	LGPQGshvvikkd...Hss...PY 30 - 47	
60 kDa chaperonin	abortus (Q5L4Q8) *	LGPQGscvikkd...Rms...PY 30 – 47 AGKVpdavvatle..Had...PY 459 - 477	
Putative conserved membrane protein	abortus (Q5L5B8)	LSLIglvlfcyke...Kipn..PY 211 - 229	-
Lipoprotein, putative	psittaci (F0T4S1)	LSLMlgvlfcyke...Kipn..PY 211 - 229	
	caviae (Q822C5)		
Putative uncharacterized protein	pecorum (F4DJN0)	LGLAwlfckke.....Kipn..PY 213 - 229	
ABC transporter, permease protein, putative	psittaci (F0T340)	FSRYirietlkq....Rnm...PY 454 - 470	Integral to membrane, transporter activity
Lipoprotein releasing system transmembrane protein	abortus (Q5L5W5)	FSQFiaypsdevykn.Rvl...PY 177 - 196	Integral to membrane, transporter activity
	psittaci (F0T459)	FSQFiaypsdevyqn.Rvl...PY 177 – 196	
ABC transporter, permease protein	caviae (Q822Z8)	FSQFvaypsdevyks.Rvl...PY 177 - 196	
Permease, putative			
Putative ABC transporter peptide periplasmic binding lipoprotein	psittaci (F0T4B9)	LGQPwhqgltkel...Kent..PY 230 - 248	Transporter activity
	abortus (Q5L5R9)		
	caviae (Q822T0)		
Peptide ABC transporter, periplasmic binding protein	caviae (Q822S8)	LGIRaidkstleitleKpi...PY 166 - 186	Transporter activity
Putative peptide ABC transport ATP-binding protein	abortus (Q5L751)	AGRLaeyapvkdif..Hnpch.PY 232 - 252	Transporter activity
Putative uncharacterized protein	psittaci (F0T3Q3)	WAFFspnsfykq...Rfst..PY 37 - 55	-
	caviae (Q823N7)	WAFFspnnfhkq...Rfst..PY 67 - 85	
	abortus (Q5L6B6)	WAFFspnsfnhkq...Rfst..PY 67 - 85	
	pecorum (F4DII8)	WAMFspnenfykq...Rfst..PY 67 - 85	
Polymorphic outer membrane protein H family	psittaci (F0T3G3)	IGLAfgglyqv....Ksk...PY 741 – 757	Outer membrane Transporter activity
Polymorphic outer membrane protein	abortus (Q5L6J8)	IGLAfgglyqv....Ksk...PY 737 - 753	
Extracellular solute-binding protein	Psittaci (F0T338)	ASKRmdlyhrfhevi.Hees..PY 593 - 613	Transporter activity
Putative peptide ABC transport system permease protein	abortus (Q5L5R6)	LAYAkglspikvilr.Hil...PY 207 - 226	Integral to membrane, transporter activity
	caviae (Q822S7)	LAYAkglspikvilr.Hil...PY 207 – 226	
Peptide ABC transporter, permease protein	pecorum (F4DJ73) *	FAYAkglpplkvilr.Hvl...PY 225 – 244	
Outer membrane protein 11	pecorum (F4DJR8)	LGFLfrkdpqlei...KpliglPY 84 – 104	autotransporter
Amino acid permease	pecorum (F4DJK1)	LSQVrpdltagiymysRkgfg.PY 67 - 89	amino acid transmembrane transporter activity

		[KR]X[KR][KR][KR]X K	
DNA-directed RNA polymerase subunit beta	psittaci (Q5MIL2)	RkKRRrK 187 - 193	transcription
	abortus (Q5L5I3)		
	caviae (Q822J1)		
Holliday junction resolvase-like protein	psittaci (F0T4B8)	RkKRKgK 122 – 128	DNA recombination / repair
Putative Holliday junction resolvase	abortus (Q5L5S0)		
Putative Holliday junction resolvase	caviae (Q822T1)	RkKRKgK 124 – 130	
Putative nucleoside triphosphate transport protein 2	abortus (Q5L6Y7)	KaKKKtK 280 - 286	Integral to membrane, ATP:ADP antiporter activity
ADP, ATP carrier protein	psittaci (F0T323)		
	caviae (Q824L9)	KaKKKtK 277 - 283	
	pecorum (F4DKP4)	KaKKKvK 272 - 278	
Putative uncharacterized protein	caviae (Q823W5)	RrRKRkK 374 – 380	-
Putative inner membrane protein	psittaci (F0T3I0)	RkKRKkK 372 - 382	
	abortus (Q5L6I7)	RdRKRrK 399 – 405 RkKRKkK 401 - 407	-
ABC transporter, periplasmic substrate-binding protein, putative	pecorum (F4DJA0)	KwKRRqK 189 - 195	

		[KR][KR]X(3)[KR][KR] K	
Hypothetical membrane associated protein	trachomatis ser D (D7DFA6) trachomatis ser D (D7DEK7) *	RKffsRKK 65 - 72	-
Putative uncharacterized protein	psittaci (F0T314)	KKkqgRKK 149 – 156	-
	abortus (Q5L6Z6)	KKkqgRKK 149 - 156	
	caviae (Q824M8)	KKkqgRKK 151 - 158	
Conserved membrane protein	abortus (Q5L504) *	RKameKRK 57 - 64	Integral to membrane
Preprotein translocase subunit	psittaci (F0T538) *		
Preprotein translocase, YajC subunit	caviae (Q821Q4) *		
Preprotein translocase, YajC subunit	pecorum (F4DJZ1) *	RKaaeKRK 56 – 63	
Efflux protein, putative	caviae (Q824X9)	KKvsyRRK 551 - 558	transmembrane transport

Tab. 2 Identified NLS carrying ptotein. The identified NLS bearing proteins are reported, In Yellow boxes are reported autotransporter proteins. Analyzed C.trachomatis pmpD NLS are hilighted in red. LEGEND: * identify proteins bearing two or more differend kind of NLS; * identify proteins bearing two or more NLS of the same Kind. Identified NLS bearing proteins are listed by mean of consensus sequence.

References

1. Blanchard and Mabey '94; Br J Clin Prac. 48: 201
2. Kuo et al. '95; Clin. Microbiol Rev. 37:423
3. Hahn and Mc Donald '98; Ann. Allergy Asthma Immunol. 81: 339
4. Kalayoglu et al. '02; JAMA 4: 288
5. Betts et al. '09; Curr Opin Microbiol. 12: 81
6. Moulder '91; Microbiol. Rev 55: 143
7. Balsara et al. '06; Infect Immun. **74**: 5602
8. Rasmussen '97; J Clin Invest. **99**: 77
9. Zhou, H., et al. '98; Nat Genet. **20**: 189
10. Pennini et al. '10; PLoS One 6: e1000995
11. Arbibe et al. '07; Nat Immunol 8: 47
12. Toyotome et al. '01; J Biol. Chem 276: 32071
13. Benabdillah et al. '04; Microb. Pathog 36: 247
14. Schornack et al. '06; J. Plant Physiol. 163: 256
15. Rossi et al. '93; Mol. Gen. Genet. 239: 345
16. Pohler et al. '95; Mol Microbiol 17: 1073
17. Peric' et al'08; Eur J cell Biol 87: 47
18. Pemberton and Paschal '05; Traffic 6: 187
19. Gorlich '99; Annu. Rev. Cell. Dev. Biol. 15: 607
20. Siomi and Dreyfuss '95; J. Cell Biol. 129:551
21. Lee et al. '06; Cell 126: 543
22. Swanson et al. '09; Infect Immun. 77: 508
23. Poon et al. '05; *Cancer Res.* 65:7059
24. Alvisi et al. '08; Biochem. 47: 13764
25. Bairoch et al. '04; Brief. Bioinform. 5: 39
26. Linding et al. '03; Structure 11:1453
27. Dautin and Bernstein '07; Annu. Rev. Microbiol. 61: 89
28. Kiselev et al. '06; PLoS One 6: e568
29. Pirbhai et al. '06; J Biol Chem 281:31495
30. Sun and Ley '08; Trends Immunol. 29: 469

Chapter II

Susceptibility of Human Placenta Derived Mesenchymal Stromal/Stem Cells to Human Herpesviruses Infection

SUSCEPTIBILITY OF HUMAN PLACENTA DERIVED MESENCHYMAL STROMAL/STEM CELLS TO HUMAN HERPESVIRUSES INFECTION

Simone Avanzi¹, Valerio Leoni², Antonella Rotola³, Francesco Alviano⁴, Liliana Solimando², Giacomo Lanzoni⁴, Laura Bonsi⁴, Dario Di Luca³, Cosetta Marchionni⁴, Gualtiero Alvisi^{2,5} and Alessandro Ripalti^{1}*

¹*Department of Oncology, Haematology and Laboratory Medicine, Operative Unit of Microbiology. A.O.-U. di Bologna Policlinico S.Orsola-Malpighi, Bologna, Italy*

²*Department of Haematology and Oncological Sciences “L.A. Seragnoli”, Microbiology Section, University of Bologna, Bologna, Italy.*

³*Department of Experimental and Diagnostic Medicine, Microbiology Section, University of Ferrara, Ferrara, Italy.*

⁴*Department of Histology, Embryology and Applied Biology, University of Bologna, Bologna, Italy.*

⁵*Department of Molecular Medicine, Microbiology Section University of Padua, Padua, Italy.*

**Alessandro Ripalti, Operative Unit of Microbiology. A.O.-U. S.Orsola-Malpighi, via Massarenti 9, 40138 Bologna, Italy.*

alessandro.ripalti@unibo.it

Tel.: +39 051 4290921

Fax.: +39 051 307397

Keywords: Human herpesviruses, Mesenchymal Stromal/Stem Cells, Placenta, Fetal Membranes, Transplantation

ABSTRACT

Fetal membrane (FM) derived mesenchymal stromal/stem cells (MSCs) are higher in number, and have greater expansion and differentiation abilities compared with those obtained from adult tissues, including bone marrow. Upon systemic administration, *ex vivo* expanded FM-MSCs preferentially home to damaged tissues promoting regenerative processes through their unique biological properties. These characteristics together with their immunoprivileged nature and immune suppressive activity, a low infection rate and young age of placenta compared to other sources of SCs make FM-MSCs an attractive target for cell-based therapy and a valuable tool in regenerative medicine, currently being evaluated in clinical trials. In the present study we investigated the permissivity of FM-MSCs to members of the *Herpesviridae* family, an issue which is relevant to their purification, propagation, conservation and therapeutic use, as well as to their potential role in the vertical transmission of viral agents to the fetus and their use in viral vector-mediated genetic modification. We present here evidence that FM-MSCs are fully permissive to infection with Herpes simplex virus 1 and 2 (HSV-1 and HSV-2), Varicella zoster virus (VZV), and Human Cytomegalovirus (HCMV), but not with Epstein-Barr virus (EBV), Human Herpesvirus-6, 7 and 8 (HHV-6, 7, 8) although these viruses are capable of entering FM-MSCs and transient, limited viral gene expression occurs. Our findings therefore strongly suggest that FM-MSCs should be screened for the presence of HSVs, HCMV and VZV before xenotransplantation. In addition, they suggest that herpesviruses may be used as viral vectors for transient gene expression in MSCs both in gene therapy applications and in the selective induction of differentiation.

INTRODUCTION

The use of nonembryonic stem cells (SCs) has opened new avenues in developmental biology and regenerative medicine. Mesenchymal stromal/cells (MSCs) constitute a heterogeneous population found first in bone marrow (BM) [2]. MSCs are easy to isolate [3], they have superior expansion potential as compared to other adult tissue-derived SCs, and are endowed with low inherent immunogenicity and the ability to modulate or /suppress immunologic responses [4]. These characteristics together with high plasticity, a tendency to migrate into damaged tissues where they orchestrate regenerative processes, and their outstanding record of safety in clinical trials make these cells prime candidates for cellular therapy. Indeed MSCs from BM or umbilical cord blood have been used in therapeutic approaches involving haematopoietic, cardiovascular, central nervous, gastrointestinal, renal, and orthopedic systems, in experimental approaches to the treatment of genetic disorders and cancer [4, 5] as well as being considered for gene therapy [6, 7].

Adult BM is a common source of MSCs for clinical use [5], however the frequency of MSCs in human adult BM is relatively low, and availability requires invasive procedures. As a consequence a quest for alternative sources of MSCs was initiated, resulting in the identification of MSCs in multiple adult and neonatal tissues such as fat, skin, cartilage, skeletal muscle, synovium, peripheral blood, dental pulp, umbilical cord, amniotic fluid and placenta [3, 8-10]. While most of these are not practical sources of MSCs for use in routine clinical applications, the human placenta at term is an alternative, ethically acceptable, abundant and easily available source of MSCs. Fetal membrane (FM) derived-MSCs are characterized by high plasticity [11-13], and are capable of differentiating into both their natural mesodermal and non mesodermal lineages [14-16], suggesting similar characteristics to BM-MSCs [17]. Amniotic membranes contribute to fetal maternal tolerance [18] and their allogeneic transplantation, or transplantation of cells derived from them, does not induce acute immune rejection even in the absence of immunosuppression [19-21]. It is not surprising therefore that FM-MSCs do not elicit allogeneic or xenogeneic immune responses, and are able to actively suppress lymphocyte proliferation [22-24]. In addition to the above mentioned therapeutic applications of MSCs, FM-MSCs are expected to be clinically used as autologous grafts for fetuses and newborns in *peripartum* tissue regeneration or for *in utero* transplantation in the case of genetic disorders without immunologic rejection by the recipient [25-27], proof of principle having already been established [28, 29]. Lastly, gene transfer in fetal blood derived MSCs with unperturbed differentiation potential has been performed [30], thus confirming their potential use in gene therapy approaches, and large scale production and manufacturing for clinical trials is being implemented [31-33].

Infections by herpesviruses are a common complication in organ transplants and during pregnancy. The human *Herpesviridae* family is composed of large, enveloped DNA viruses with close structural similarity and includes the Herpes simplex virus types 1 and 2 (HSV-1 and 2), Varicella zoster virus (VZV), Epstein Barr virus (EBV), Human Cytomegalovirus (HCMV), as well as Human Herpesvirus (HHV) types 6, 7 and 8. All members of the family replicate in the nucleus of the infected cell after activating a coordinated cascade of mRNA synthesis, that drives gene transcription in three temporal classes: immediate early (IE), early (E) and late (L). These viruses all share the ability to establish latency and reactivate at a later time [34]. Allogeneic SC transplantation is often complicated by reactivation of herpesviruses, which are also a cause of morbidity and mortality in patients undergoing solid organ transplants [35]. The presence of selected viruses in MSCs from donor BM has been investigated [36], human fetal

BM-MSCs support transcription of HHV-8 immediate early LANA messenger [39], while adult BM-MSCs have proved permissive to HSV-1 and HCMV [37, 38], raising concerns regarding their possible transmission to transplant recipients. Furthermore several herpesviruses can establish congenital infections: structural fetal abnormalities can result from intrauterine infection and transmission of the infection during pregnancy or at the time of delivery and can result in severe neonatal disease [40].

In the present study we aimed to investigate the susceptibility of FM-MSCs to all members of the human *Herpesviridae* family. As a whole, studying FM-MSCs permissivity to herpesviruses is relevant to: (i) the transplantation field, in that infection of plastic cells used in regenerative medicine pre or post transplant could interfere with their differentiation process and compromise the engraftment, inducing cytopathology or death of the implanted cells; (ii) vertical transmission of infections, a setting where the mechanism of virus spread through the placenta is vastly unexplored, and the consequences of congenital infection have been poorly characterized for most members of the family, especially in relation to stillbirth and post partum long term sequelae; (iii) gene therapy and viral vector mediated genetic modification of SCs.

Our results clearly show that FM-MSCs are fully permissive to infection by HSV-1/2, VZV and HCMV. On the other hand, despite EBV, HHV-6, HHV-7 and HHV-8 being capable of entering FM-MSCs, only limited gene expression occurs, resulting in a nonproductive infection.

MATERIALS AND METHODS

Antibodies.

(i) **Surface antigens analysis:** mouse monoclonal antibodies anti-CD21 (clone LB21) and anti-CD35 (clone E11), were purchased from Antibodies-Online (Antibodies-Online INC., Atlanta, USA); anti-CD19 (clone HD37) obtained from DAKO (DAKO, Glostrup, Denmark); anti-CD45 (clone 2D1, APC), CD11c (S-HCL-3, PE), CD14 (MΦP9, PE), CD31 (WM-59, PE), CD36 (NL07, FITC), CD90 (5E10, FITC), CD59 (P282-H19, PE), CD184 (12G5, PE) and CD166 (3A6, PE) were purchased from BD (Becton-Dickinson and Company, New York, USA); anti-CD106, (1G1b1-PE) were obtained from Southern Biotechnology (Southern Biotechnology Associates, Birmingham, USA); HLA-DR (Tu36, FITC), CD 105 (SN6, PE), CD44 (MEM 85, FITC), HLA-ABC (Tu149, FITC), CD80 (MEM-233, PE), CD29 (MEM101A, PE) were purchased from Caltag Laboratories (Caltag Laboratories, Bangkok, Thailand), CD31 (CBL468F, FITC); CD34 (HPCA-2, Bdis, APC or PE-Cy7) were obtained from Cymbus Biotechnology (Cymbus Biotechnology, Chandlers Ford, UK).

(ii) **Indirect ImmunoFluorescence (IIF):** monoclonal mouse antibodies raised against HSV-1 and HSV-2 gD, clone HDI was a kind gift from Gabriella Campadelli-Fiume (University of Bologna, Bologna, Italy). HCMV anti-IE clone E13 and anti-pp65, clone 1C3 were purchased from ARGENE SA (Varilhes, France). Anti-UL44 clone 10D8 was obtained from Virusys (Virusys corp., Taneytown, USA). Fluorescein (clone sc2010) or Rodhamine (clone sc2092) conjugated secondary antibodies were purchased from Santa Cruz (Santa Cruz Biotechnology Inc., Santa Cruz, USA). All antibodies were diluted in PBS, according to the manufacturer's instructions.

Cells. Human embryo lung derived fibroblasts (HEL) were cultured in minimal essential medium (MEM, PAA, Pasching, China), supplemented with 10% fetal calf serum (FCS) and 100 U/ml penicillin and 100 µg/ml streptomycin (P/S). Human umbilical vascular endothelial cells (HUVECs), were maintained in endothelial cell growth medium supplemented with 5% FCS and growth factors (EGM-MV single aliquots, both from BioWhittaker, Cambrex Bio Science, Walkersville, MD). Vero, and BHK cells were maintained in Dulbecco's modified Eagle's medium (DMEM, EuroClone, Milan, Italy) supplemented with 5% FCS, and P/S. Both T lymphoid J-Jhan and Sup T1 cell lines were grown in suspension in RPMI 1640 (EuroClone, Milan, Italy) supplemented with 10% FCS.

FM-MSCs isolation, ex vivo culture and characterization. According to the policy approved by the local Ethical Committee (S.Orsola-Malpighi University Hospital, Bologna, Italy), all tissue samples were obtained after informed written consent. FM-MSCs were isolated as described previously [41]. Briefly, term placentas from healthy donor mothers undergoing caesarean sections were rapidly transferred to the laboratory, rinsed in Phosphate Buffered Saline (PBS) containing penicillin and streptomycin (200 U/ml penicillin, 200 µg/ml streptomycin) and processed immediately. Pieces of fetal membrane were minced and subjected to 15-minute digestion with 0.25% trypsin-EDTA solution. The supernatant was discarded and the tissue underwent a second digestion with 0.25% trypsin-EDTA solution, 10 U/ml DNaseI and 0.1% collagenase IV solution in DMEM (all from Sigma-Aldrich, St. Louis, USA). Larger pieces of tissue were allowed to settle under gravity for 5 min at 37°C, while each supernatant was transferred to a fresh tube, neutralized with Fetal Bovine Serum (FBS, Biochrom, Berlin, Germany), then spun down at 400 g for 10 min. Each pellet was resuspended in 5 ml of DMEM containing 20% FBS, and P/S 100 µg/ml. Cells were seeded in 25 cm² flasks and grown at 37°C in 5% CO₂. Non-adherent cells were removed after one week and the medium (with 10 % of FBS) was subsequently changed every four days. When the culture reached 90% confluence, cells were passaged using 0.25% Trypsin-EDTA.

Surface antigens were characterized as described elsewhere [41]. Briefly, trypsinized cells were resuspended in cell culture medium (10% FBS), incubated on ice for 10 min, washed with Dulbecco's phosphate-buffered saline (PBS, pH 7.2, 10% FBS, Invitrogen, Carlsbad, USA), and labeled with 1 $\mu\text{g}/10^6$ cells FITC-conjugated antibodies for 40 min at 4°C in the dark.

Infection. Target cells were seeded either on 24- or 6-well culture plates at a cell density of 10^4 cells/cm². When required, cells were layered onto coverslips in 24-well plates, grown until they reached approximately 80% confluence and infected as described below.

Viruses. An HSV-1 clinical isolate at low passage from our diagnostic laboratory was used in most experiments. Alternatively, we used a fully replication competent derivative of the HSV-1 KOS strain, K26GFP, expressing VP26 as a fusion with the green fluorescent protein (GFP), which leads to incorporation of GFP in the capsid [42], thus allowing easy direct identification of infected cells with an inverted fluorescence microscope. The low passage HSV-2 (G) laboratory strain was used in this study. Virus stocks were prepared and titrated in Vero (HSV-1 and K26GFP) and BHK (HSV-2) cells. Briefly, supernatants (SNs) from infected cells approaching 100% cytopathic effect (cpe) were collected and subjected to three rounds of 20 sec sonication with a Bandelin Sonoplus sonicator, Model HD-2200 with an SH213G probe (Bandelin Electronic GmbH, Berlin, Germany) allowing a 1 min cooling-off period between each sonication. Cellular debris were pelleted by centrifugation at 800 xg for 20 min at 4°C, and the SNs were aliquoted and stored at -70°C. Titration was performed by thawing frozen aliquots, which were sequentially tenfold diluted and used in triplicate to infect Vero or BHK cells layered onto coverslips in 24-well plates for 1 h at 37°C, in a 5% CO₂ atmosphere. The inoculum was then removed, cells were rinsed with 1 ml PBS and fed with fresh medium. 6 h post infection (p.i.) a neutralizing antibody was added to the medium in order to avoid extracellular spread of progeny virions. Plaques were counted three days post p.i. in wells where the dilution of the viral stock allowed less than 10 plaques to appear, and the correspondent titre was calculated. FM-MSCs were infected with virus stocks at a multiplicity of infection (moi) varying from 1-2.5 plaque forming units (pfu)/cell. Infection was performed for 1 h at 37°C in a humidified CO₂ incubator, then the inoculum was removed and fresh medium added.

For VZV, a clinical isolate at low passage from our diagnostic laboratory was used in all experiments. Virus stocks were prepared in Vero cells by infection with cell-associated virus at a ratio 1:4, infected to uninfected cells in 75 cm² plastic flasks, harvesting infected cells when cpe was > 80 % with a cell scraper in their medium. Cell debris were subjected to three rounds of sonication as for HSV, except that cell debris were not separated from the medium after sonication, but were directly aliquoted and stored at -70°C. Rough titrations were performed by infection of Vero cells with serial dilutions of the frozen stocks in 24-well plates, overlaying a 0.6% agarose containing culture medium 24 h p.i. and counting plaques at three days p.i. FM-MSCs were infected with VZV virus stocks at a multiplicity of infection (moi) < 1 pfu/cell. Infection was performed for 1 h at 37°C in a humidified CO₂ incubator, then the inoculum was removed and fresh medium added.

As far as HCMV is concerned, we used the AD169 laboratory strain and the low passage endotheliotropic TB40 strain. Viral stocks of TB40 were prepared in HUVECs, while AD169 was prepared in HEL. Viral stocks were titrated in HEL as described elsewhere [43]. FM-MSCs and HEL were infected with HCMV viral stocks with an moi varying between 0.1 to 2.5 pfu/cell. Infection was performed for 1 h at 37°C in a humidified CO₂ incubator, then the inoculum was removed and fresh medium added.

HHV-6 viral stocks were obtained as previously described [44]; HHV-6 variant A (strain U1102) was grown in the T lymphoid J-Jhan cell line [45]; HHV-6 variant B (strain CV) was grown in the Sup T1 cell line [46]. A single inoculum was prepared and used for all infection experiments to avoid variability in the efficiency of infection due to differences in inocula. Briefly, cell-free viral inocula were obtained by pelleting a total of 500 ml of cell cultures (at a concentration of 1×10^6 cells/ml) infected with HHV-6A or HHV-6B exhibiting complete cpe. The cells were lysed by 3 cycles of rapid freezing and thawing followed by sonication (3 cycles of 5 sec at medium power with 10 second intervals in water bath sonicator). Cleared cellular content was added to culture supernatant and virions were collected by centrifugation at 20000 xg at 4°C. Virus particles were purified by density centrifugation using Optiprep self-forming gradients (Sentinel, Milan, Italy) at 5.8×10^4 xg for 3.5 h at 4°C. Collected virions were suspended in PBS containing 1% bovine serum albumin, fractioned in aliquots of 100 µl and stored at -80°C until use. Prior to use, virus stocks were treated with DNase-I and RNase A, to eliminate free viral nucleic acids eventually present in the preparation. HHV-6 viral stocks were quantitated by real time PCR as already described [47] using primers which amplify both HHV-6 variants. The standard curve was generated by amplification of a plasmid containing the targeted HHV6 sequences. The method had a 6-log dynamic range and a sensitivity of 20 copies/ml. FM-MSCs were infected with HHV-6A or B with a moi of 50 genomes/cell. Infection was performed for 2 h at 37°C in a humidified CO₂ incubator, then the inoculum was removed and fresh medium added.

HHV7 Strain CZ [48] was grown in the Sup T1 cell line. Cell free viral inoculum was obtained as previously described [46]. Briefly, 1 l of HHV-7 infected cell cultures exhibiting complete cpe was centrifuged at 900 xg, resuspended in 2 ml of FCS supplemented with RNase (50 µg/ml; Boehringer Ingelheim, Milan, Italy) and disrupted by four cycles of freezing in liquid nitrogen and thawing at 37°C. The resulting inoculum was completely free of living cells, as checked by microscopic observation and cultivation, and was also analyzed by reverse transcription PCR (both for β -actin and for a panel of viral mRNAs) to ensure that RNA was completely absent. 1×10^6 FM-MSCs were infected with 400 µl of the resulting inoculum. Infection was performed for 2 h at 37°C in a humidified CO₂ incubator, the inoculum was removed and fresh medium added.

The B95-8 cell line was used to generate infectious EBV virus. Cells were propagated in complete RPMI medium at 37°C and an atmosphere of 5% CO₂, and fed 2 times per week, maintaining cell density between $0.5\text{--}1.0 \times 10^6$ /ml. On day 0 12-O-tetradecanoylphorbol-13-acetate (TPA) was added to a final concentration of 25 ng/ml together with sodium butyrate to a final concentration of 4 mM. The cells were harvested on day 5. Cells that were not induced were also harvested on the fifth day. Supernatant from both TPA induced and non induced B95-8 cell lines was used to infect FM-MSCs after clearance at low speed of cellular debris and subsequent filtration through a 0.8 mm-pore filter to remove any remaining debris. Titration of inoculi was performed by determining the number of genomes present in serial dilutions of the final suspension with a real time PCR, targeting the exon 4/5 of the terminal protein gene with primers EB-1 5' AACATTGGCAGCAGGTAAGC and EB-2 5' ACTTAC-CAAGTGTCCATAGGAGC, and producing a 182 pb amplicon [49].

FM-MSCS infection was performed at 37°C in a humidified CO₂ incubator for 2 h, then the inoculum was removed and replaced with fresh medium. Mock infections was performed with virus free medium in the same conditions. HEL cells were infected or mock infected in the same conditions in parallel.

When infected cells were cultured in 24 well chambers with coverslips at the bottom of each well, cells were analysed by IIF at 24, 48 and 72 h p.i., and every 3 days thereafter up to 21 days p.i.

BC-3 cells, a KSHV/HHV-8 positive primary effusion lymphoma (PEL) – derived cell line, were used as a source of HHV-8 and virus was purified by a modification of the protocol described by Caselli and co-workers [50]. Cells were propagated in RPMI supplemented with 20% fetal bovine serum at 37°C and an atmosphere of 5% CO₂, fed every 2-3 days for 3-4 weeks until a minimal amount of 2×10^8 cells was obtained. Cells were then treated with TPA at 20ng/ml (Sigma-Aldrich, St. Louis, USA) for 60 h, subsequently centrifuged and the supernatant was collected. The cell pellet was lysed by three freeze-thaw cycles, centrifuged to eliminate cell debris, pooled with the supernatant, filtered through a 0.45-µm-pore-size filter, and concentrated at 4°C overnight with 7% polyethylene glycol 8000 (Sigma-Aldrich, St. Louis, USA). The suspension was centrifuged at 1.5×10^4 xg for 2 h at 4°C and the pellet was resuspended in a small amount of phosphate-buffer saline (PBS) and purified through a 25% sucrose cushion (2.65×10^4 xg for 3 h in a Beckman JSW-13-1 rotor). The pellet was resuspended in PBS and layered onto a 40-70% sucrose gradient and centrifuged at 2.65×10^4 xg for 1 h at 4°C in a Beckman JSW-13-1 rotor. The visible band was collected, dialyzed O/N against PBS, aliquoted and then stored at -80°C. Serial dilutions of this stock suspension were titrated by determining the number of genomes, with a real time PCR targeting the LANA gene, using primers HH8-1 5'TTGGGAAAGGATGGAAGACG and HH8-2 5'AGTCCCCAGGACCTTGTTT, which determine an amplification product of 346 bp [51].

Nucleic Acid Purification. At specific time-points after infection cells were harvested by trypsinization, washed twice in cold sterile PBS, collected by centrifugation at 700 xg and stored at -80°C until processed. For nucleic acid extraction and purification cells were thawed, resuspended in 100µl of cold PBS and total nucleic acid were extracted with the NucliSENS easyMAG automatic extractor (bioMérieux sa, Marcy l'Etoile, France), following the manufacturer's instructions. Purified material was used both for DNA and RNA analysis.

Polymerase chain reaction (PCR) for detection of viral DNA in donor FM-MSCs. 2 multiplex PCR assays were performed, HERP-1 for HSV-1 and 2, VZV, EBV and HERP-2 for HCMV, HHV-6, HHV-7 and HHV-8. Primers included were previously described by others [49, 51, 52].

PCR analysis for detection of HHV-6A and B, and HHV-7 in infected FM-MSCs. To determine which samples harbored viral genomic DNA sequences, 1 µg of DNA (corresponding to 1.5×10^5 cells) was analyzed by nested PCR by using primers amplifying the U42 gene both for HHV-6A and B, and HHV-7. All primers were derived from the published HHV-6 and 7 sequences [53, 54]. The primer sequences as well as the PCR conditions were as previously described [45, 46, 55, 56].

Reverse Transcription-PCR (RT-PCR) for the detection of viral transcripts in infected FM-MSC. DNA enzymatic degradation was performed by means of DNase I Amplification Grade (Invitrogen, Carlsbad, CA, USA). RT-PCR was performed with the SuperScript™ III CellsDirect cDNA Synthesis System (Invitrogen, Carlsbad, CA, USA), and specific sequences were PCR amplified in an automated thermal cycler, using the Perfect Taq Plus amplification kit (5 PRIME GmbH, Hamburg, Deutschland) according to manufacturer instructions. For detection of EBV, the three splicing form of the EBNA-1 transcript was amplified using the set of oligonucleotides previously published by Chen and co-workers [57], following their working conditions. BZLF1 first round RT-PCR was performed using 1 µl of cDNA in a 50 µl reaction volume and 1.5 mM MgCl₂ final concentration, using oligonucleotides BZLF1-OUT-Fw 5'AGCAGACATTGGTGTTCAC and BZLF1-OUT-Rv 5'ACATCTGCTTCAA-CAGGAGG. Amplification reactions were carried out incubating samples at 95°C for 2 min then a 35 round cycle constituting incubations at 95°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec, followed by a final incubation at

72°C for 10 min. BZLF1 nested RT-PCR was performed with the same amplification conditions in combination with primers BZLF1-IN-Fw 5'ACGCACGGAAACCACAAC and BZLF1-IN-Rv 5'GCGCAGCCTGTCATTTT-CAG. RT-PCR for the detection of HHV-6A and HHV-6B transcripts was performed as described elsewhere [45, 58], while HHV-7 specific transcripts were analyzed as described in [46]. In the case of HHV8, ORF73 transcript was amplified using oligonucleotides ORF73-OUT-Fw 5'GAAGTGGATTACCCTGTTGTTAGC and ORF73-OUT-Rv 5'AGTCCCCAGGACCTTGGTTT and 2.5 µl of cDNA as template. ORF 73 semi-nested PCR was performed using oligos ORF73-OUT-Fw and ORF73-IN-Rv. 5'TATCTCAGGCCTTCCAGTTT. ORF50 transcript was amplified with oligonucleotides ORF50-OUT-Fw 5'GCCCTCTGCCTTTTGGTT and ORF50-OUT-Rv 5'GATGATGCTGACGGTGTG. Semi-nested PCR was performed with oligonucleotides ORF50-IN-Fw 5'GCAAGGTCAGTGGACTGTC and ORF50-OUT-Rv. All amplification reactions were carried out in a 50 µl reaction volume, with 1.5 mM MgCl₂ final concentration, using 2.5 µl of DNA as template. A common amplification protocol was used, with the following conditions: 95°C for 2 min, 40 x (95°C for 30 sec, 60°C for 30 sec and 72°C for 60 sec), 72°C for 10 min final amplification.

Fluorescence Microscopy. Following fixation with paraformaldehyde 4% (w/v), cells were incubated with primary antibody for one h at 37°C in a humidified chamber. Three wash steps were subsequently performed in PBS, five min each at room temperature (RT). A secondary antibody, labeled with a fluorophore and raised against the species producing the primary antibody, was then incubated with the cells, in the same conditions as for the primary antibody. Antibody dilutions varied and were generally in compliance with manufacturers instructions. When human serum samples from our collection were used, dilution was generally 1:50. Cells expressing GFP-VP26 were incubated with DAPI (1 µg/mL) for 4 min at RT to stain cell nuclei, after fixation with paraformaldehyde 4% (w/v). Samples were washed with PBS, mounted on coverslips in PBS/glycerol 50% (v/v), and imaged using a Nikon Eclipse E600 microscope equipped with a Nikon DXN1200 digital camera and a Nikon Plan Fluor 40x objective (Nikon, Tokyo, Japan). Live cells grown on Willcodishes (Willcowsells, Amsterdam, The Netherlands) expressing GFP-VP26 fusion protein were imaged using a Nikon Eclipse E600 inverted microscope equipped with a Nikon DXN1200 digital camera and a Nikon Plan Fluor 40x objective (Nikon, Tokyo, Japan), as previously [42, 59, 60].

RESULTS

Herpesvirus genomes are not found in term placenta derived MSCs from healthy (seropositive) individuals

FM-MSCs obtained from term placentas of 6 donors were analysed for the presence of herpesviruses genomes by means of 3 multiplex PCR assays developed in house, HERP-1 for HSV-1 and 2, and VZV, HERP-2 for HCMV, HHV-6, and HHV-7 and HERP-3 for EBV and HHV-8 [61]. No traces of any herpesvirus genome were found in any of the six populations of FM-MSCs used in this study (data not shown).

Permissivity of FM-MSCs to herpesviruses

To assess the infection efficiency of viral stocks, cell lines permissive to HHVs infection were infected with the same moi used for FM-MSCs and time-points were collected.

All RNA samples were analysed by PCR without retrotranscription to check the total absence of DNA contamination. No contaminant DNA was present in any sample (see figures, RT-). Furthermore the efficiency of retrotranscription was monitored by amplification of the human β-actin cDNA, after a 1:100 dilution.

Alpha herpesviruses

HSV-1 and HSV-2

MSCs originating from bone marrow have been shown to be permissive to HSV-1 [37], while heparan sulfate has been suggested as a major determinant for virus entry in adipose tissue derived MSC [63]. To investigate the degree of permissivity of FM-MSCs to HSV-1 and HSV-2, monolayers of FM-MSCs were infected with a clinical isolate of HSV-1, or the HSV-2 laboratory strain G. When infected cells were observed under an optical inverted microscope, typical HSV-induced cpe became apparent from 16-24 h p.i at high moi (Fig. 1A), whereas cytopathic effects were observable only at later times after infection at low moi. Furthermore, IIF analysis allowed the specific detection of the L glycoprotein gD (Fig. 1B). Observation of cells infected at high moi from 18 h p.i. allowed the identification of polykaryons (Fig. 1B, bottom panels), whose formation requires L viral proteins gB, gD, gH and gL, that can be synthesised only after IE and E transcripts have been translated and DNA replication has been initiated. To verify that infectious progeny virus was released in the SN of infected cells, SNs were collected from infected cells at 2, 18 and 24 h p.i. and infectious virus was titrated on Vero or BHK cells in the case of HSV-1 or HSV-2, respectively. Our results show that infectious progeny virions arise from infected FM-MSCs, reaching titers of up to 10^7 p.f.u./ml (Fig. 1C). Interestingly, when FM-MSCs and control Vero cells were infected with the K26GFP recombinant HSV-1 virus, the VP26-GFP fusion protein was readily detectable in both cell lines, similarly accumulating in the host cell nucleus by 20 h p.i. (Supplementary Fig. 1). In conclusion, FM-MSCs can be regarded as fully permissive for both HSV-1 and HSV-2 viruses.

VZV

Viral genomic DNA has been PCR amplified from BM-MSCs isolated from two osteoarthritis patients in a sample of 18 patients [38]. However, susceptibility of MSC to VZV *in vitro* has never been addressed to date. We assayed the permissivity of FM-MSCs to VZV by infecting monolayers of FM-MSCs, or HEL as a positive control, with cell-associated VZV (moi < 1). When infected cells were observed under an inverted optical microscope, cytopathic foci became clearly visible from 4 days p.i. both in HEL and in FM-MSCs (Fig. 2A, left panels). At 6 days p.i. signs of cell fusion, a hallmark of VZV infection started to appear (Fig. 2A, middle panels). The effect was more pronounced in FM-MSCs, where larger polykaryocytes were observable as compared to HEL. This was even more evident at 9 days p.i. when cytopathic effect was generalized in both cell populations. The presence of infectious progeny virions was assayed on SNs from infected cells. SNs were collected 6 days p.i. and processed for nucleic acid extraction and purification. PCR analysis was used to identify VZV genomic DNA in the extracted material. As shown in Fig. 2B, VZV specific amplicons were obtained from samples corresponding to SNs of infected HEL and FM-MSCs. Our results clearly show that FM-MSCs are fully permissive to VZV.

Beta herpesviruses

HCMV

BM-MSC have been shown to be susceptible to the AD169 strain [37] and to the clinical isolate TB40 [38] of HCMV. In a similar way we tested permissivity of FM-MSCs to both AD169 and TB40 viruses. To this end, we infected FM-MSCs, and HEL as positive controls, with both viruses at an moi of ≈ 1 pfu/cell. At this moi, a majority of cells in an infected HEL monolayer will allow progression of the viral replication cycle; the percentage of cells hosting a full viral replication cycle in a cell population with unknown permissivity to HCMV compared to

that of a population of HEL infected in parallel with a comparable inoculum, reflects the permissivity of the cell type. Cells were fixed at different time points thereafter. IIF assays were then performed with antisera directed to specific virus encoded products. In order to determine if representative members of all temporal classes of viral genes were expressed in these cells, we analysed the expression of IE, E, and L genes at 24-96 h p.i. in both cell populations. As shown in Fig. 3A, viral encoded genes of all temporal classes of expression were expressed both in infected HEL and FM-MSCs. To analyze permissivity of FM-MSCs at the single cell level, counts were made of IIF positive cells for each antigen studied, at the indicated time points, and expressed as a percentage of the whole population. When the expression of the major IE gene UL123 was analysed upon infection with TB40 from 24 to 96 h p.i. 82% of the cell population was found positive at 24 h, 90% at 48 h, 93% at 72 h and 95% at 96 h p.i. When the expression of the E gene ORFUL44 was analysed at 24 h and 48 h p.i., 58.5% and 72.6% of FM-MSCs stained positive respectively, while staining of the matrix protein expressed by the L gene ORFUL99 at 72 h and 96 h p.i. resulted in 87% and 96% positivity. When HEL cells were infected with TB40, or FM-MSCs and HEL were infected with a comparable inoculum of the AD169, we obtained similar results. As a whole these data suggest that FM-MSCs are fully permissive to HCMV, and that a replication cycle takes between 72 and 96 h to complete when cells are infected at a moi of 1 pfu/cell.

To evaluate the kinetics of replication of HCMV in FM-MSCs at high moi as compared to HEL cells, and verify that infectious progeny virions are released by infected FM-MSCs, we infected both cell populations at a moi of 10 pfu/cell with the TB40 strain. At this moi, the viral replication cycle is synchronized in all cells of the monolayer, and in HEL cells at 72h pi progeny virions will appear in culture supernatants. The medium of infected cell monolayers was therefore collected after 3 days, centrifuged to eliminate cell debris and frozen, while cells were fed with fresh medium. SNs were then collected in the same way with a 3 day time interval until the monolayers were completely lysed by the virus. A titration of all collected supernatants was then performed on HEL. As shown in Fig. 3B virus yield in both cell populations reached a peak at 6 days p.i. and decreased gradually. HEL appeared more resistant to the virus than FM-MSCs, the latter monolayer being completely lysed 2-3 days earlier. This accounts for a 2 log lower virus yield at 18 days p.i. in FM-MSCs compared to HEL cultures. Taken together these results indicate that FM-MSCs are fully permissive to HCMV.

HHV-6

FM-MSCs were infected with a cell free inoculum of HHV-6A(U1102) or HHV-6B(CV). The inocula for the two viruses were standardized, quantitating the viral genome by real-time PCR. Infection was performed using a moi of 50 viral genomes/cell. The cultures were analysed at different times (1, 3, 7, 14, 21 and 28 days p.i.). Infected cultures had a morphology similar to uninfected cultures and no cpe was detected (data not shown). DNA was extracted at all time points and analyzed by single round PCR and nested PCR for the presence of HHV-6A and HHV-6B DNA, with a PCR reaction specific for U42. For both variants, viral DNA was present at all time points, up to 21 days p.i. after single step PCR, and up to 28 days p.i. by nested PCR (Fig. 4A). These results suggested that the amount of intracellular viral DNA decreased over time, and therefore real-time quantitation of viral DNA was carried out. The results of these experiments showed that the amount of intracellular HHV-6 DNA decreased from approximately 5×10^7 genome equivalents in 1.5×10^4 cells 24 h p.i. to 5×10^2 at 28 days p.i., and that both variants exhibited the same behaviour (Fig. 4B). Therefore, it is plausible to postulate that viral DNA persisted in infected cells with low levels of viral replication.

To determine the replicative state of HHV-6 in infected FM-MSCs, viral transcription was analysed by RT-PCR for the presence of IE (U42) and L (U22) viral lytic transcripts. The analysis also included U94/*rep*, a latency-associated transcript that is expressed at low levels during productive infection [45]. The sensitivity of PCR reactions was similar for all genes, detecting 10^3 target molecules consistently after the first round of PCR and 10-50 target molecules after nested PCR. Cells infected with HHV-6A showed the presence of U42 and U22 transcripts by first round PCR until 14 days p.i., and the transcript of U94 was detected till 21 days p.i. All three transcripts persisted as long as 28 days p.i., as shown by nested PCR (Fig. 4C). These results suggest that HHV-6A establishes a low level productive infection, characterized by scarce viral replication and decreasing amounts of viral products (both DNA and RNA) over time.

In the case of cells infected with HHV-6B, single round PCR failed to show the presence of viral transcripts, and nested PCR yielded positive results at all time points only for U42 and U94. The transcript of U22 was detected only during the first 3 days of infection (Fig. 4C). These results suggest that viral replication compatible with lytic production took place only for a few days, and that afterwards HHV-6B replication was restricted to the early phase of infection, with persistence of IE transcripts (U42 and U94), and the absence of late mRNAs (U22).

Low grade HHV-6 infection and consequently low levels of viral replication may account for the lack of evident cytopathic effect in FM-MSC, which is clearly seen as large ballooning cells in primary human mononuclear cells or in selected T cell lines (JJhan).

HHV-7

FM-MSCs were infected with a cell free inoculum of HHV-7 strain CZ. The cultures were analysed at different times (1, 3, 7, 14, 21 and 28 days p.i.). Infected cultures had a morphology similar to uninfected cultures and no cytopathic effect was detected (data not shown). DNA was extracted at all time points and analysed by single round PCR, with primers and conditions specific for U42 of HHV-7. Viral DNA was detected at all time points, without the need for nested amplification (Fig. 5A). No precise quantitation was performed, however the low intensity of amplification bands at 21 and 28 days p.i. suggests that the amount of viral DNA decreased over time. Analysis of viral transcripts was carried out by RT-PCR, in order to monitor expression of U42, U16/17 and U89/90 mRNAs. All these genes belong to the IE transcriptional class and are expressed at high levels throughout viral replication, starting immediately after infection [46]. Only U42 mRNA was detected at all time points, with a faint band after single step PCR. U16/17 and U89/90 mRNA were not detected, even after nested PCR (Fig. 5B). On the whole, these results show that HHV-7 infects FM-MSCs cells, and that the virus persists without fully replicating, and expressing only some immediate-early functions [63].

Gamma herpesviruses

EBV

To assess the susceptibility of FM-MSCs to EBV infection, supernatant from a TPA induced B98-5 cell line was used to infect FM-MSCs and EBV susceptible HEK cells at a genome/cell ratio of 100. SN from non induced cells and serum free medium were used to establish mock infected controls. The presence of the virus did not cause any morphological change either at the cell monolayer or at the single cell level as observed by light microscopy.

Amplification of DNA extracts from infected and mock infected cells and culture SNs from day 1 to 28 p.i. was performed, to investigate the presence of viral genomes in infected FM-MSCs. As shown in Fig. 6A genomic DNA

was amplified at up to 28 days p.i., with generally very weak signals. However, when the same samples were subjected to a nested PCR analysis, specific DNA was amplified at all time points, suggesting that low levels of viral genomic DNA persists in the infected population. No DNA was amplified in any sample originating from infected or control cell SNs (data not shown). When virus specific transcripts were analysed in genomic extracts of infected cells by RT-PCR, as described in materials and methods, no viral mRNAs were detected at any time point investigated (Fig. 6B). Cytometric analysis of FM-MSCs revealed CD21 surface expression. These results taken together strongly suggest that EBV virions can enter FM-MSCs but transcription is not allowed while viral DNA persists for at least 4 weeks p.i. Since the cell monolayer is confluent by day 3/4 after infection, it is unlikely that failure to identify traces of viral transcripts or genome amplification is due to a dilution effect within a proliferating population of cells.

HHV-8

MSC from fetal BM can be infected latently in vitro by HHV-8, and virus persistence and expression of LANA/ORF70 has been documented for over 9 weeks in those cells [39]. To assess the ability of HHV-8 to infect FM-MSC, frozen stocks of purified virus were diluted in serum free medium and added to FM-MSCs at a moi of 50 genomes/cell. Serum free medium was used to establish mock-infected controls. HUVEC cells were infected in parallel as positive controls. Infected cells remained unperturbed by the presence of the virus at all times p.i. as observed microscopically (data not shown). The presence of intracellular and extracellular viral DNA was assessed by PCR amplification of nucleic acid extracts obtained at days 1 to 28, from cell monolayer extracts and extracellular fluid. Qualitative and quantitative analyses were performed on both extracts as described in materials and methods. Viral DNA was readily detected at all time points tested (Fig. 7A).

Analysis of virus encoded transcripts was performed on RNA extracted from mock infected and infected cells from 1 to 28 days p.i., and analysed by RT-PCR. ORF50 and ORF73 mRNAs were retro-transcribed and resulting cDNAs were amplified with the specified primers as described in materials and methods. As shown in Fig. 7B we did not detect any of the two transcripts at any time point analysed. When the infection was performed at high moi ($>10^3$ genomes/cell) both ORF50 and ORF73 generated transcripts were detectable at 3 and 6 days p.i. (Fig. 7C), but not at later time points (data not shown). Coexpression of a marker of latent infection (ORF73) and of the lytic-switch protein (ORF50) suggests a general, albeit short lived activation of immediate early viral promoters, while the following cascade of genomic transcription would not be supported by the cellular transcription machinery, nor would the transcriptional activity of ORF50 be allowed to proceed.

When quantitative analysis of virus genome copy number was implemented in high genome/cell ratio HHV-8 infected FM-MSC, 1 genome per 25 cells was detected at 3 days p.i., 1 genome per 288 cells was found at 6 days p.i., while at 9 days p.i. the amount of DNA was under the detection limit. Taken together these results suggest that HHV-8 is able to penetrate FM-MSCs and PCR detectable amounts of its genome persist for at least 4 weeks p.i., however limited gene expression is allowed only following high moi and for less than one week.

Passage number and herpesvirus susceptibility

We have tested the ability of FM-MSCs to sustain infection of the eight members of the human *Herpesviridae* family starting at the 5th passage up to the 20th passage. No significant differences in infectability, virus yield, mRNA transcription or genome number and persistence were found in cells at the different expansion stages analysed, confirming the already described phenotypic stability of these cells throughout extensive expansion (data not shown).

DISCUSSION

The present work addresses for the first time the susceptibility of FM-MSCs to all members of the *Herpesviridae* family. We show here that FM-MSCs are susceptible to HSV-1, HSV-2, VZV and HCMV infection in vitro. Susceptibility remains unaltered from the 5th to the 20th passage. Our results show that HHV-6A persists in FM-MSCs cells with low levels of complete viral expression, and therefore with potential for production of infectious virus. On the other hand, HHV-6B and HHV-7 show a more restricted expression, with persistence of viral DNA and continuous transcription of some immediate-early genes. Even if late genes are not expressed, it is not possible to exclude the possibility that these viruses might undergo full replication under specific conditions of the host cell. HHV-8 persists for at least 4 weeks in infected FM-MSCs, while specific transcripts are detectable only upon high moi, but for not more than 6 days pi. Our data are consistent with the observation by Parsons and co-workers that expression of ORF73 is supported in fetal BM MSC upon infection at high genome/cell ratio [39]. In our experimental settings duration of ORF73 transcription in infected cells was shorter, very likely due to lower doses of inoculi.

Ethical issues restrict the use of antenatal tissues as a source of SCs for scientific research and clinical applications. As an alternative one can rely upon the isolation of SCs from available fetal tissue samples, as with fetal bioptic or blood sampling and terminated pregnancies, or term placenta and postnatally retrieved FM, which yield mainly SCs with mesenchymal cell surface markers and morphology, MSCs. Amnion and chorion in particular constitute a valuable source of MSCs [11, 13, 14].

Herpesviruses are widely spread in the population, with seroprevalence rates ranging from a minimum of 30% for HCMV and 60% for HSV-1 in some Western developed countries, to over 90% for VZV, HHV-6 and 7 and EBV in the rest of the world [64].

Our study, showing that FM-MSCs are fully permissive to members of the *Herpesviridae* family has implications in three medical research fields. Firstly in regenerative medicine, because herpesvirus infections remain a major challenge both in solid organ and hematopoietic stem cell transplantation, two settings in which the use of FM-MSCs is being implemented with growing expectation. Awareness of the possible contamination of isolated FM-MSCs with viruses suggests the need to screen isolated cells for the presence of viral genomes, in order to avoid virus transmission to transplant recipients. Furthermore, the infection of transplanted tissues originating from the expansion of these cells could have adverse consequences on the outcome of the engraftment. Considerazioni su good manufacturing practices.

A second field of interest is that of vertical transmission of herpesviruses. Primary infection with HSV has been associated with spontaneous abortion, premature labor and intrauterine growth retardation, while primary VZV infection may result in stillbirth or cause congenital varicella, a syndrome associated with a significant mortality rate [65]. HCMV is the leading cause of congenital viral infection and the most common infectious agent of congenital malformations in developed countries. Viruses and other difficult to culture organisms have been postulated as the aetiology of a number of obstetric and paediatric conditions of unknown cause, including stillbirth. Among herpesviruses, HCMV, HSV, and VZV may cause intrauterine deaths [66]. Infection of FM-MSCs with a member of the *Herpesviridae* family could have negative consequences on pregnancy and potentially the fetus at all times

of gestation. MSC from vascularized FMs could be a first target for replication and dissemination of virus to non vascularized FMs and to fetal circulation, both serving as a reservoir of actively replicating virus, and interfering with the immune response of resident T cells. Since FM-MSCs have been shown to enhance fetal repair [67] and couple immunosuppressant abilities to antimicrobial effector function [68], their depletion by viral infection could further complicate viral damage to the fetus. Furthermore herpesviruses are known to cause chromosome breaks in infected cells [69], that could represent a teratogenic threat to the fetus considering that MSCs traffic between placenta membranes and the fetus [70], especially in very early developmental stages. Furthermore, since the placenta is not a barrier to maternal MSC trafficking, and BM derived MSC have been successfully infected *in vitro* with HSV-1 and HCMV, it is conceivable that vertical transmission could occur via this straightforward route. The severity of depletion or functional impairment of FM-MSCs during early gestation could depend upon maternal immune competence towards HCMV which has been shown to occur for cytotrophoblasts [71] and could disrupt placenta implantation or cause abnormal placenta development, resulting in early failure. It is important to note here that our placenta-derived cells grow in cell culture conditions. The degree of their representativeness with respect to the original cell population constituting the stromal layer of mesenchymal tissue within fetal membranes has been considered with the intrinsic limitation due to an *in vitro* model. However, considering the high amount of MSC that can be isolated from a tiny area of fetal membranes, the number of passages needed to characterize the cell population is very limited (3-5), while our experiments were carried out at passage 6, when we would not expect major shifts in MSCs biochemical and biological characteristics would have occurred. Finally our work is relevant to the field of virus mediated gene therapy applied to SCs. The importance of reprogramming FM-MSCs functions of interest or orienting their differentiation potential via gene transfer has been established and has a broad consensus [72, 73].

Because of their propensity to migrate to sites of injury and their ability to expand rapidly, FM-MSCs have been considered as potential gene transfer vehicles to deliver therapeutic genes [6, 7, 74]. On the other hand, gene transfer to MSC is facilitated by using viral vectors [75-77]. Our results contribute to better predict how genetically modified versions of the different herpesviruses may be used to this end. Using members of the *Herpesviridae* family for gene transfer can have several advantages: long-term transgene expression, repeat vector dosing with minimal immune response due to the absence of helper viruses during viral packaging, and simultaneous delivery of large and multiple transgenes [78]. Hybrid HSV/EBV vectors have been used to enhance stability of transgene expression [79]. Their different behaviour in terms of genome stability and differential genome expression in infected cells may be advantageous in distinct applications as diverse as transient expression for the induction of differentiation or long term expression of therapeutic genes.

Considerazioni su good manufacturing practices e sicurezza negli impieghi clinici (Umbilical Cord Blood-Derived Mesenchymal Stem Cells Inhibit, But Adipose Tissue-Derived Mesenchymal Stem Cells Promote, Glioblastoma Multiforme Proliferation).

SUPPORTING INFORMATION

Supplementary Fig. 1 Expression of a virus unrelated protein by a herpes simplex based viral vector. Recombinant GFP expressed by the replication defective vector TOZGFP-HSV1, carrying deletions in the immediate early genes ICP4, ICP22 and ICP27. GFP localizes in infected FM-MSCs and Vero with a comparable pattern.

ACKNOWLEDGEMENTS

The authors thank Gabriella Campadelli-Fiume (Bologna, Italy) and Prashant Desai (Baltimore, USA) for kindly providing monoclonal mouse antibodies raised against HSV-1 and HSV-2 gD, Elisabetta Caselli and Monica Galvan for kindly providing HHV-6 and HHV-7 viral stocks, and Roberto Manservigi and Peggy Marconi for the generous gift of the replication deficient virus TOZGFP-HSV1. This work was supported by Assegni di Ricerca, Università di Bologna (LS); Fondi Ateneo per la Ricerca, Università di Ferrara (ARo and DDL); Associazione Italiana per la Ricerca sui Virus (SA and VL).

AUTHORS CONTRIBUTION

Collected and assembled data: SA, VL, LS, FA, ARo; Analysed and interpreted data: GA, SA, CM, DDL, VL, ARi, ARo; Contributed reagents/materials/analysis tools: GL, FA, ARo, CM; Conceived and designed the experiments: ARi; Wrote the paper: ARi, ARo; revised the paper: SA, FA, GA, DDL, GL, VL, ARi, ARo.

REFERENCES

1. Strioga, M., et al., *Same or Not the Same? Comparison of Adipose Tissue-Derived Versus Bone Marrow-Derived Mesenchymal Stem and Stromal Cells*. Stem Cells Dev, 2012.
2. Minguell, J.J., A. Erices, and P. Conget, *Mesenchymal stem cells*. Exp Biol Med (Maywood), 2001. **226**(6): p. 507-20.
3. Phinney, D.G. and D.J. Prockop, *Concise review: mesenchymal stem/multipotent stromal cells: the state of transdifferentiation and modes of tissue repair--current views*. Stem Cells, 2007. **25**(11): p. 2896-902.
4. Hematti, P., *Role of mesenchymal stromal cells in solid organ transplantation*. Transplant Rev (Orlando), 2008. **22**(4): p. 262-73.
5. Brooke, G., et al., *Therapeutic applications of mesenchymal stromal cells*. Semin Cell Dev Biol, 2007. **18**(6): p. 846-58.
6. Hwang, D.H., S.R. Jeong, and B.G. Kim, *Gene transfer mediated by stem cell grafts to treat CNS injury*. Expert Opin Biol Ther, 2011. **11**(12): p. 1599-610.
7. Hai, C., et al., *Application of mesenchymal stem cells as a vehicle to deliver replication-competent adenovirus for treating malignant glioma*. Chin J Cancer, 2012. **31**(5): p. 233-40.
8. Kolf, C.M., E. Cho, and R.S. Tuan, *Mesenchymal stromal cells. Biology of adult mesenchymal stem cells: regulation of niche, self-renewal and differentiation*. Arthritis Res Ther, 2007. **9**(1): p. 204.
9. He, Q., C. Wan, and G. Li, *Concise review: multipotent mesenchymal stromal cells in blood*. Stem Cells, 2007. **25**(1): p. 69-77.
10. van den Berk, L.C., C.G. Figdor, and R. Torensma, *Mesenchymal stromal cells: tissue engineers and immune response modulators*. Arch Immunol Ther Exp (Warsz), 2008. **56**(5): p. 325-9.
11. In 't Anker, P.S., et al., *Isolation of mesenchymal stem cells of fetal or maternal origin from human placenta*. Stem Cells, 2004. **22**(7): p. 1338-45.
12. Campagnoli, C., et al., *Identification of mesenchymal stem/progenitor cells in human first-trimester fetal blood, liver, and bone marrow*. Blood, 2001. **98**(8): p. 2396-402.
13. Fukuchi, Y., et al., *Human placenta-derived cells have mesenchymal stem/progenitor cell potential*. Stem Cells, 2004. **22**(5): p. 649-58.
14. Parolini, O., et al., *Concise review: isolation and characterization of cells from human term placenta: outcome of the first international Workshop on Placenta Derived Stem Cells*. Stem Cells, 2008. **26**(2): p. 300-11.
15. Bossolasco, P., et al., *Molecular and phenotypic characterization of human amniotic fluid cells and their differentiation potential*. Cell Res, 2006. **16**(4): p. 329-36.
16. Portmann-Lanz, C.B., et al., *Placental mesenchymal stem cells as potential autologous graft for pre- and perinatal neuroregeneration*. Am J Obstet Gynecol, 2006. **194**(3): p. 664-73.
17. Yen, B.L., et al., *Isolation of multipotent cells from human term placenta*. Stem Cells, 2005. **23**(1): p. 3-9.
18. Mellor, A.L. and D.H. Munn, *Immunology at the maternal-fetal interface: lessons for T cell tolerance and suppression*. Annu Rev Immunol, 2000. **18**: p. 367-91.
19. Ward, D.J. and J.P. Bennett, *The long-term results of the use of human amnion in the treatment of leg ulcers*. Br J Plast Surg, 1984. **37**(2): p. 191-3.
20. Faulk, W.P., et al., *Human amnion as an adjunct in wound healing*. Lancet, 1980. **1**(8179): p. 1156-8.
21. Subrahmanyam, M., *Amniotic membrane as a cover for microskin grafts*. Br J Plast Surg, 1995. **48**(7): p. 477-8.
22. Wolbank, S., et al., *Dose-dependent immunomodulatory effect of human stem cells from amniotic membrane: a comparison with human mesenchymal stem cells from adipose tissue*. Tissue Eng, 2007. **13**(6): p. 1173-83.
23. Bailo, M., et al., *Engraftment potential of human amnion and chorion cells derived from term placenta*. Transplantation, 2004. **78**(10): p. 1439-48.
24. Li, H., et al., *Immunosuppressive factors secreted by human amniotic epithelial cells*. Invest Ophthalmol Vis Sci, 2005. **46**(3): p. 900-7.
25. Surbek, D.V., W. Holzgreve, and K.H. Nicolaides, *Haematopoietic stem cell transplantation and gene therapy in the fetus: ready for clinical use? Hum Reprod Update*, 2001. **7**(1): p. 85-91.
26. Liechty, K.W., et al., *Human mesenchymal stem cells engraft and demonstrate site-specific differentiation after in utero transplantation in sheep*. Nat Med, 2000. **6**(11): p. 1282-6.
27. Chan, J., et al., *Widespread distribution and muscle differentiation of human fetal mesenchymal stem cells after intrauterine transplantation in dystrophic mdx mouse*. Stem Cells, 2007. **25**(4): p. 875-84.
28. Tsuda H., et al., *Allogenic fetal membrane-derived mesenchymal stem cells contribute to renal repair in*

- experimental glomerulonephritis. *Am J Physiol Renal Physiol*, 2010. **299**(5): F1004-13.
29. Ohshima, M., et al., *Systemic transplantation of allogenic fetal membrane-derived mesenchymal stem cells suppresses Th1 and Th17 T cell responses in experimental autoimmune myocarditis*. *J Mol Cell Cardiol*, 2012. **53**(3): p. 420-8.
30. Chan, J., et al., *Human fetal mesenchymal stem cells as vehicles for gene delivery*. *Stem Cells*, 2005. **23**(1): p. 93-102.
31. Steigman, S.A. and D.O. Fauza, *Isolation of mesenchymal stem cells from amniotic fluid and placenta*. *Curr Protoc Stem Cell Biol*, 2007. **Chapter 1**: p. Unit 1E 2.
32. Brooke, G., et al., *Manufacturing of human placenta-derived mesenchymal stem cells for clinical trials*. *Br J Haematol*, 2009. **144**(4): p. 571-9.
33. Yu, Y., et al., *Ex Vitro Expansion of Human Placenta-Derived Mesenchymal Stem Cells in Stirred Bioreactor*. *Appl Biochem Biotechnol*, 2009.
34. *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*, ed. C.-F.G. Arvin A, Mocarski E, Moore PS, Roizman B, Whitley R, Yamanishi K. 2007, Cambridge: Cambridge University Press.
35. Ljungman P., *Risk assessment in haematopoietic stem cell transplantation: viral status*. *Best Pract Res Clin Haematol.*, 2007. **20**(2): p. 209-17.
36. Rollin, R., et al., *Human parvovirus B19, varicella zoster virus, and human herpesvirus-6 in mesenchymal stem cells of patients with osteoarthritis: analysis with quantitative real-time polymerase chain reaction*. *Osteoarthritis Cartilage*, 2007. **15**(4): p. 475-8.
37. Sundin, M., et al., *Mesenchymal stem cells are susceptible to human herpesviruses, but viral DNA cannot be detected in the healthy seropositive individual*. *Bone Marrow Transplant*, 2006. **37**(11): p. 1051-9.
38. Smirnov, S.V., et al., *Bone-marrow-derived mesenchymal stem cells as a target for cytomegalovirus infection: implications for hematopoiesis, self-renewal and differentiation potential*. *Virology*, 2007. **360**(1): p. 6-16.
39. Parsons, C.H., B. Szomju, and D.H. Kedes, *Susceptibility of human fetal mesenchymal stem cells to Kaposi sarcoma-associated herpesvirus*. *Blood*, 2004. **104**(9): p. 2736-8.
40. Schleiss, M.R., *Vertically transmitted herpesvirus infections*. *Herpes*, 2003. **10**(1): p. 4-11.
41. Ventura, C., et al., *Hyaluronan mixed esters of butyric and retinoic Acid drive cardiac and endothelial fate in term placenta human mesenchymal stem cells and enhance cardiac repair in infarcted rat hearts*. *J Biol Chem*, 2007. **282**(19): p. 14243-52.
42. Desai, P. and S. Person, *Incorporation of the green fluorescent protein into the herpes simplex virus type 1 capsid*. *J Virol*, 1998. **72**(9): p. 7563-8.
43. Ripalti, A., et al., *Cytomegalovirus-mediated induction of antisense mRNA expression to UL44 inhibits virus replication in an astrocytoma cell line: identification of an essential gene*. *J Virol*, 1995. **69**(4): p. 2047-57.
44. Caselli, E., et al., *Human herpesvirus 8 enhances human immunodeficiency virus replication in acutely infected cells and induces reactivation in latently infected cell*. *Blood*, 2012. **106**: p. 2790-2797.
45. Rotola, A., et al., *U94 of human herpesvirus 6 is expressed in latently infected peripheral blood mononuclear cells and blocks viral gene expression in transformed lymphocytes in culture*. *Proc Natl Acad Sci U S A*, 1998. **95**(23): p. 13911-6.
46. Menegazzi, P., et al., *Temporal mapping of transcripts in human herpesvirus-7*. *J Gen Virol*, 1999. **80** (Pt 10): p. 2705-12.
47. Caruso, A., et al., *U94 of human herpesvirus 6 inhibits in vitro angiogenesis and lymphangiogenesis*. *Proc Natl Acad Sci U S A*, 2009. **106**(48): p. 20446-51.
48. Portolani, M., et al., *Isolation of human herpesvirus 7 from an infant with febrile syndrome*. *J Med Virol*, 1995. **45**(3): p. 282-3.
49. Markoulatos, P., et al., *Laboratory diagnosis of common herpesvirus infections of the central nervous system by a multiplex PCR assay*. *J Clin Microbiol*, 2001. **39**(12): p. 4426-32.
50. Caselli, E., et al., *Human herpesvirus 8 acute infection of endothelial cells induces monocyte chemoattractant protein 1-dependent capillary-like structure formation: role of the IKK/NF-kappaB pathway*. *Blood*, 2007. **109**(7): p. 2718-26.
51. Ashshi, A.M., P.E. Klapper, and R.J. Cooper, *Detection of human cytomegalovirus, human herpesvirus type 6 and human herpesvirus type 7 in urine specimens by multiplex PCR*. *J Infect*, 2003. **47**(1): p. 59-64.
52. Fujimuro, M., et al., *Multiplex PCR-based DNA array for simultaneous detection of three human herpesviruses, EVB, CMV and KSHV*. *Exp Mol Pathol*, 2006. **80**(2): p. 124-31.
53. Gompels, U.A., et al., *The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution*. *Virology*, 1995. **209**(1): p. 29-51.
54. Nicholas, J., *Determination and analysis of the complete nucleotide sequence of human herpesvirus*. *J Vi-*

- rol, 1996. **70**(9): p. 5975-89.
55. Mirandola, P., et al., *Temporal mapping of transcripts in herpesvirus 6 variants*. J Virol, 1998. **72**(5): p. 3837-44.
56. Gonelli, A., et al., *Human herpesvirus 7 is latent in gastric mucosa*. J Med Virol, 2001. **63**(4): p. 277-83.
57. Chen, F., et al., *A subpopulation of normal B cells latently infected with Epstein-Barr virus resembles Burkitt lymphoma cells in expressing EBNA-1 but not EBNA-2 or LMP1*. J Virol, 1995. **69**(6): p. 3752-8.
58. French, C., et al., *Novel, nonconsensus cellular splicing regulates expression of a gene encoding a chemokine-like protein that shows high variation and is specific for human herpesvirus 6*. Virology, 1999. **262**(1): p. 139-51.
59. Alvisi, G., et al., *Nuclear import of HSV-1 DNA polymerase processivity factor UL42 is mediated by a C-terminally located bipartite nuclear localization signal*. Biochemistry, 2008. **47**(52): p. 13764-77.
60. Alvisi, G., et al., *The flexible loop of the human cytomegalovirus DNA polymerase processivity factor ppUL44 is required for efficient DNA binding and replication in cells*. J Virol, 2009. **83**(18): p. 9567-76.
61. Degli Esposti Merli, M., *Design and production of multiplex PCR Molecular Standards for the diagnostic detection of human herpesviruses*, in Dipartimento di Ematologia e Scienze Oncologiche L&A Seragnoli. 2010, University of Bologna: Bologna. p. 40.
62. Choudhary, S., et al., *Herpes simplex virus type-1 (HSV-1) entry into human mesenchymal stem cells is heavily dependent on heparan sulfate*. J Biomed Biotechnol, 2011. **2011**: p. 264350.
63. Chow, S.S., et al., *Correlates of placental infection with cytomegalovirus, parvovirus B19 or human herpes virus 7*. J Med Virol, 2006. **78**(6): p. 747-56.
64. Looker, K.J., G.P. Garnett, and G.P. Schmid, *An estimate of the global prevalence and incidence of herpes simplex virus type 2 infection*. Bull World Health Organ, 2008. **86**(10): p. 805-12, A.
65. Enright, A.M. and C.G. Prober, *Neonatal herpes infection: diagnosis, treatment and prevention*. Semin Neonatol, 2002. **7**(4): p. 283-91.
66. Rawlinson, W.D., et al., *Viruses and other infections in stillbirth: what is the evidence and what should we be doing?* Pathology, 2008. **40**(2): p. 149-60.
67. Klein, J.D., et al., *Amniotic mesenchymal stem cells enhance normal fetal wound healing*. Stem Cells Dev, 2011. **20**(6): p. 969-76.
68. Meisel, R., et al., *Human but not murine multipotent mesenchymal stromal cells exhibit broad-spectrum antimicrobial effector function mediated by indoleamine 2,3-dioxygenase*. Leukemia, 2011. **25**(4): p. 648-54.
69. Fortunato, E.A. and D.H. Spector, *Viral induction of site-specific chromosome damage*. Rev Med Virol, 2003. **13**(1): p. 21-37.
70. Barinaga, M., *Cells exchanged during pregnancy live on*. Science, 2002. **296**: p. 2169-72.
71. Maidji, E., et al., *Developmental regulation of human cytomegalovirus receptors in cytotrophoblasts correlates with distinct replication sites in the placenta*. J Virol, 2007. **81**(9): p. 4701-12.
72. Fierro, F.A., et al., *Effects on proliferation and differentiation of multipotent bone marrow stromal cells engineered to express growth factors for combined cell and gene therapy*. Stem Cells, 2011. **29**(11): p. 1727-37.
73. Munoz Ruiz, M. and J.R. Regueiro, *New tools in regenerative medicine: gene therapy*. Adv Exp Med Biol. **741**: p. 254-75.
74. Rath, P., et al., *Stem cells as vectors to deliver HSV/tk gene therapy for malignant gliomas*. Curr Stem Cell Res Ther, 2009. **4**(1): p. 44-9.
75. Pincha, M., et al., *Lentiviral vectors for induction of self-differentiation and conditional ablation of dendritic cells*. Gene Ther, 2011. **18**(8): p. 750-64.
76. Neumann, A.J., et al., *Enhanced Adenovirus Transduction of hMSCs Using 3D Hydrogel Cell Carriers*. Mol Biotechnol, 2012.
77. Cho, Y.H., et al., *Enhancement of MSC adhesion and therapeutic efficiency in ischemic heart using lentivirus delivery with periostin*. Biomaterials. **33**(5): p. 1376-85.
78. Burton, E.A., et al., *Multiple applications for replication-defective herpes simplex virus vectors*. Stem Cells, 2001. **19**(5): p. 358-77.
79. Sia, K.C., et al., *Hybrid herpes simplex virus/Epstein-Barr virus amplicon viral vectors confer enhanced transgene expression in primary human tumors and human bone marrow-derived mesenchymal stem cells*. J Gene Med, 2010. **12**(10): p. 848-58.

FIGURES AND LEGENDS

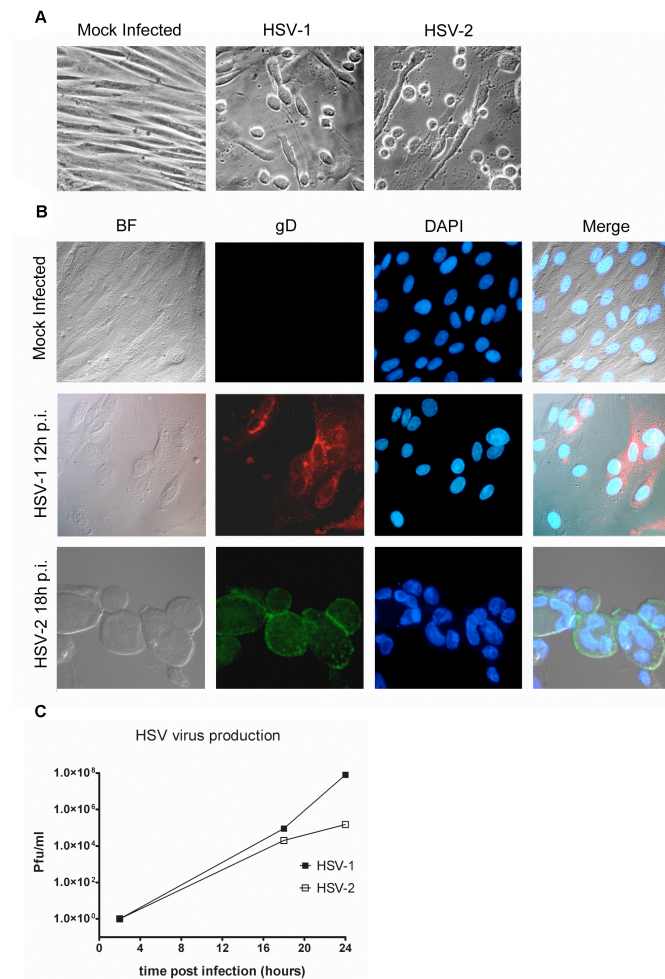


Figure 1. FM-MSCs are susceptible to HSV-1 and HSV-2 infection. (A) FM-MSCs were either mock infected (*left panel*) or infected with > 1 p.f.u./cell of the indicated viruses before being analyzed with an inverted optical microscope as described in the Materials and Methods section. The typical cytopathic effect shown is observed starting 12-18 h p.i.; (B) FM-MSCs were either mock infected (*top panels*) or infected with > 1 p.f.u./cell of the indicated viruses before being fixed and processed for IFF using a mAb directed against the late glycoprotein gD to detect infected cells, and DAPI to visualize cell nuclei. The bright field (BF) is shown in the left panels, and a merged image of all channels is shown in the right panels. Formation of typical syncytia in infected cells is evident starting at 18 h p.i., as shown in the HSV-2 bottom panels; (C) SNs of infected FM-MSCs were collected at the indicated times p. i. and used to reinfect Vero (for HSV-1) or BHK (for HSV-2) cells to quantify infectious virion yields in the extracellular fluid of FM-MSCs by means of plaque assays as described in the Materials and Methods section.

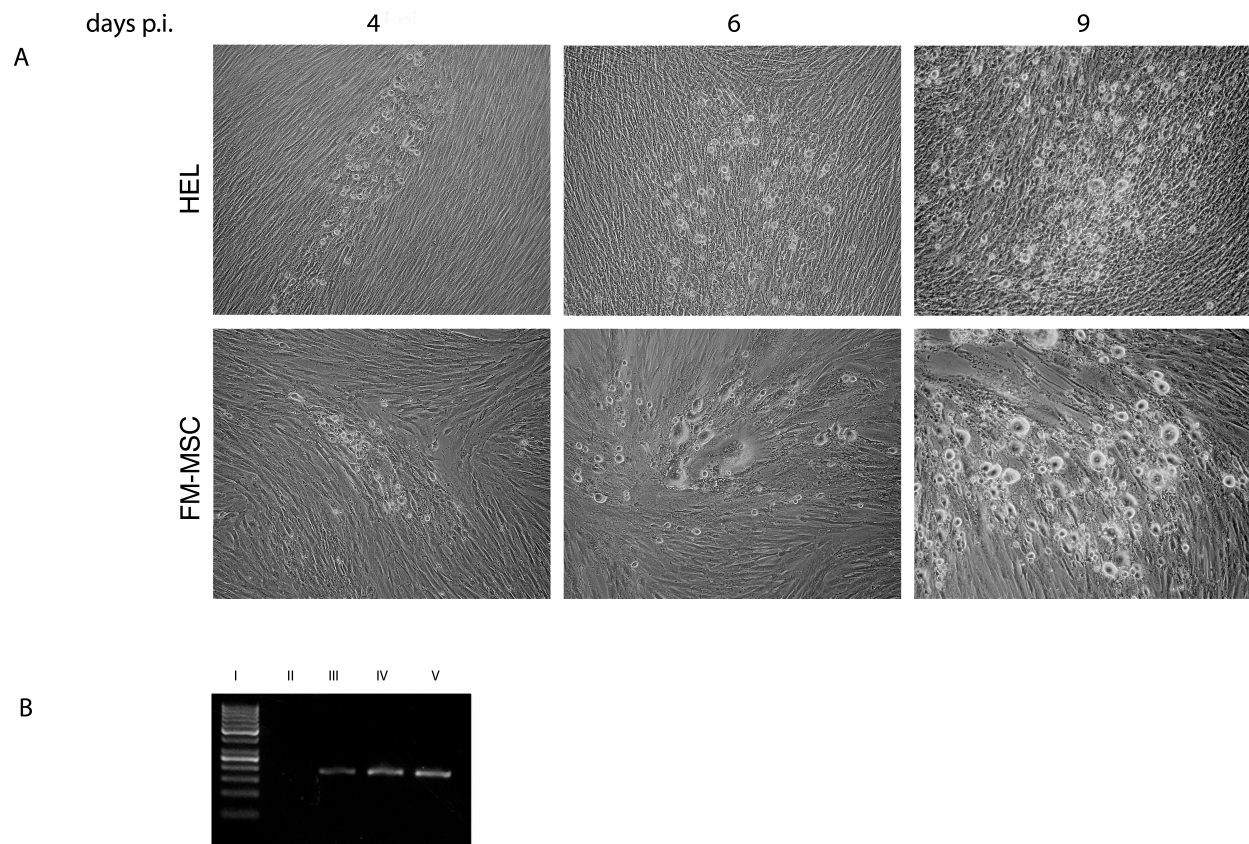


Figure 2. FM-MSCs are fully permissive to VZV infection. (A) HEL or FM-MSCs were infected with VZV at a moi < 1 pfu/cell, as described in the Materials and Methods section. At the indicated time p.i. cells were analyzed by optical microscopy; (B) to determine if the observed cytopathic effect observed in (A) was indeed due to VZV infection, VZV DNA was PCR amplified from extracellular fluid (including cellular debris): I. DNA standard, II. Negative control (SN from mock-infected FM-MSCs), III. Supernatant from VZV infected HEL, IV. Supernatant from VZV infected FM-MSCs, V. positive control (DNA extracted from VZV inoculi).

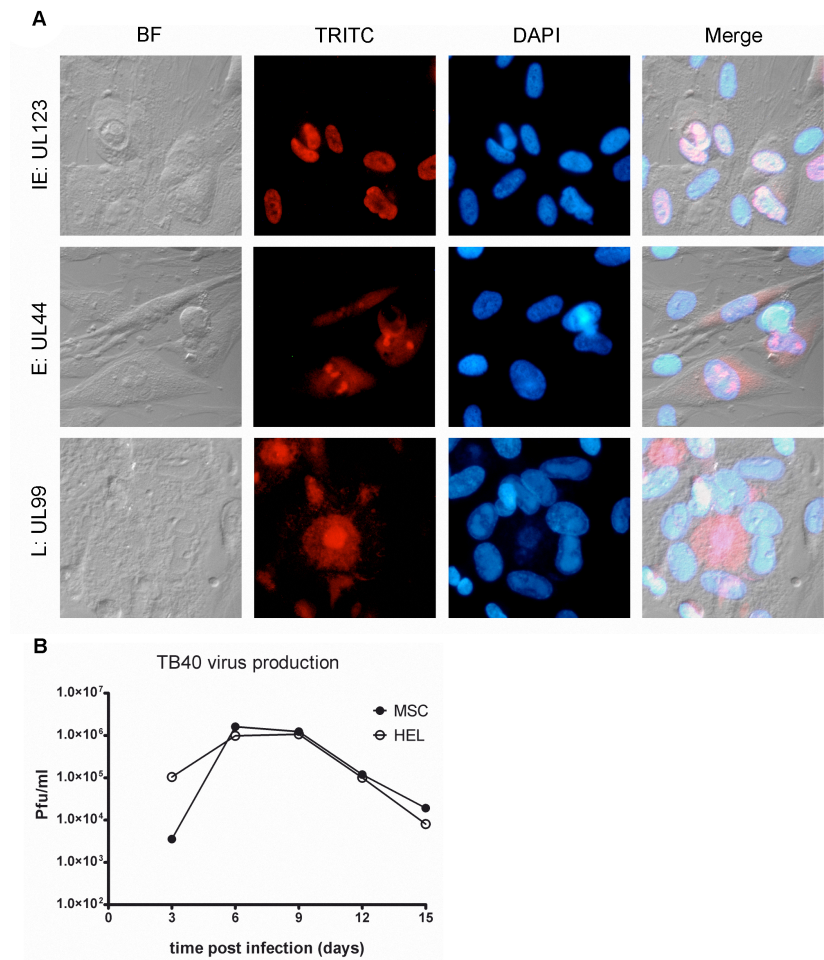


Figure 3. FM-MSCs are fully permissive to HCMV infection. (A) FM-MSCs cells were infected with HCMV (strain TB40) at a moi of 1 pfu/cell and fixed at 24 (top panels), 48 (middle panels) or 96 (bottom panels) h post p.i. before being fixed and processed for IIF to stain HCMV temporal class representative antigens using IE (UL123), E (UL44) or L (UL99) specific antibodies, as indicated. The bright field (BF), the specific antibody signal (TRITC) and the cell nuclei (DAPI) are shown, with merged images presented in the right panels; (B) SNs from infected HEL cells or FM-MSCs were harvested at the indicated times p.i. and used to determine the titers of produced HCMV infectious virions by performing plaque assays on HEL cells as described in the Materials and Methods section.

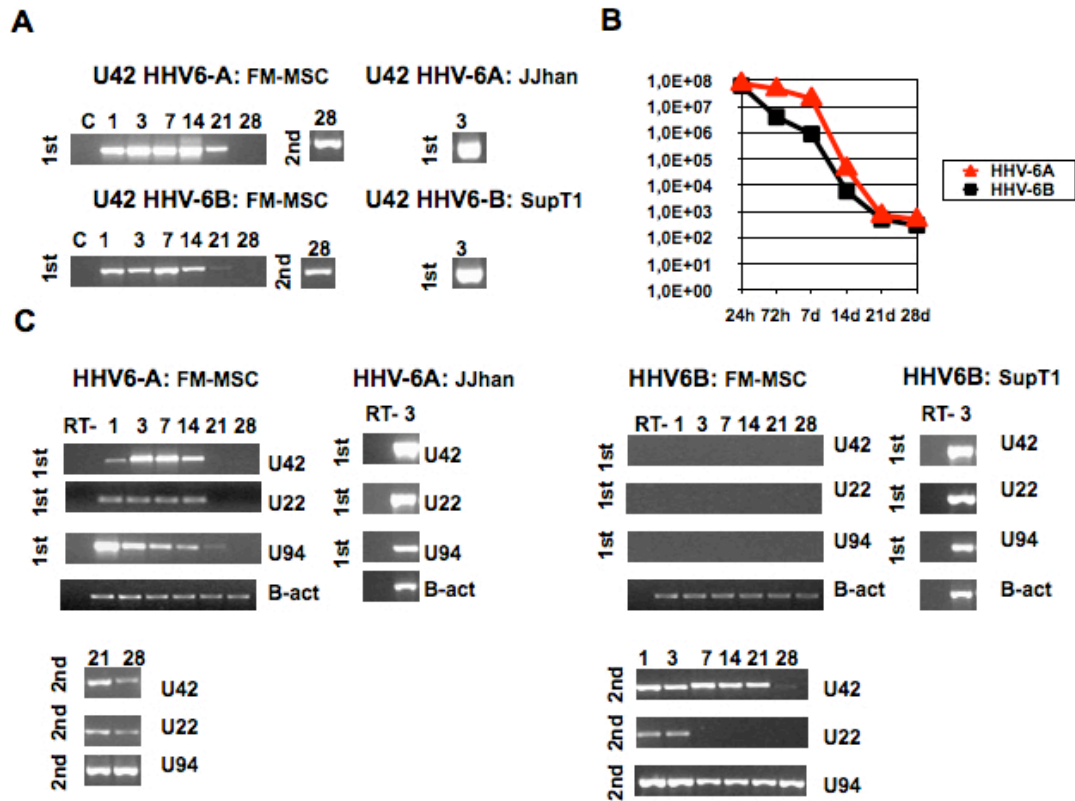


Figure 4. FM-MSCs are susceptible to HHV-6 A/B variants. FM-MSCs were infected with HHV-6 (variant A and variant B) at a moi of 50 genomes/cell and nucleic acids extracted at the indicated times p.i. As positive control an equivalent number of cells permissive to HHV-6 productive infection (J-Jhan for HHV-6A and SupT1 for HHV-6B) were infected with the same viral stocks used for infecting the stem cells. A) single step PCR (1st) and nested (2nd) detection of HHV-6A and B genomes both in FM-MSC cells and J-Jhan and SupT1 cell lines. B) Real Time PCR quantification of HHV-6A and -B genomes. C) single step PCR (1st) and nested (2nd) RT-PCR for U42 (Early), U22 (Late) and U94 (latency associated) viral transcripts both in FM-MSC cells and J-Jhan and SupT1 cell lines. The β -actin retrotranscription-positive control is also shown. C= mock infected FM-MSCs cells. RT- = PCR amplification of RNA samples without retrotranscription.

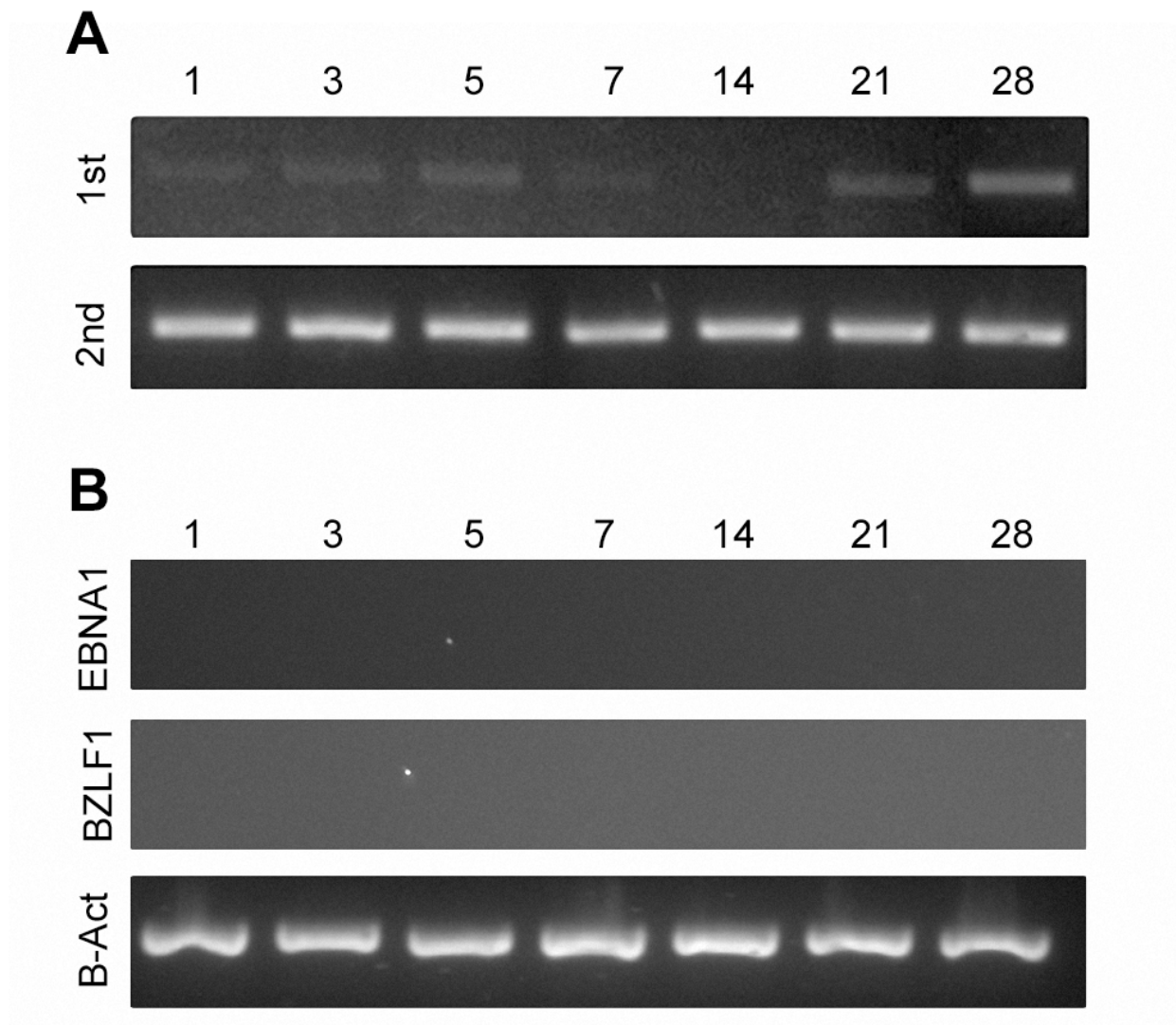


Figure 6. FM-MSCs are not permissive to EBV productive infection. (A) nested qualitative PCR performed on EBV infected FM-MSCs nucleic acid extracts: in the first amplification round viral DNA is amplified at all time points, although at barely non detectable levels. Second round amplification confirmed viral genomic persistence in infected cells; (B) nested RT-PCR against latency transcript EBNA-1 or lytic cycle associated BZLF1 transcript, on nucleic acids extracted from EBV-infected FM-MSCs failed to detect EBV transcriptional activity both in first and second round amplification (data not shown).

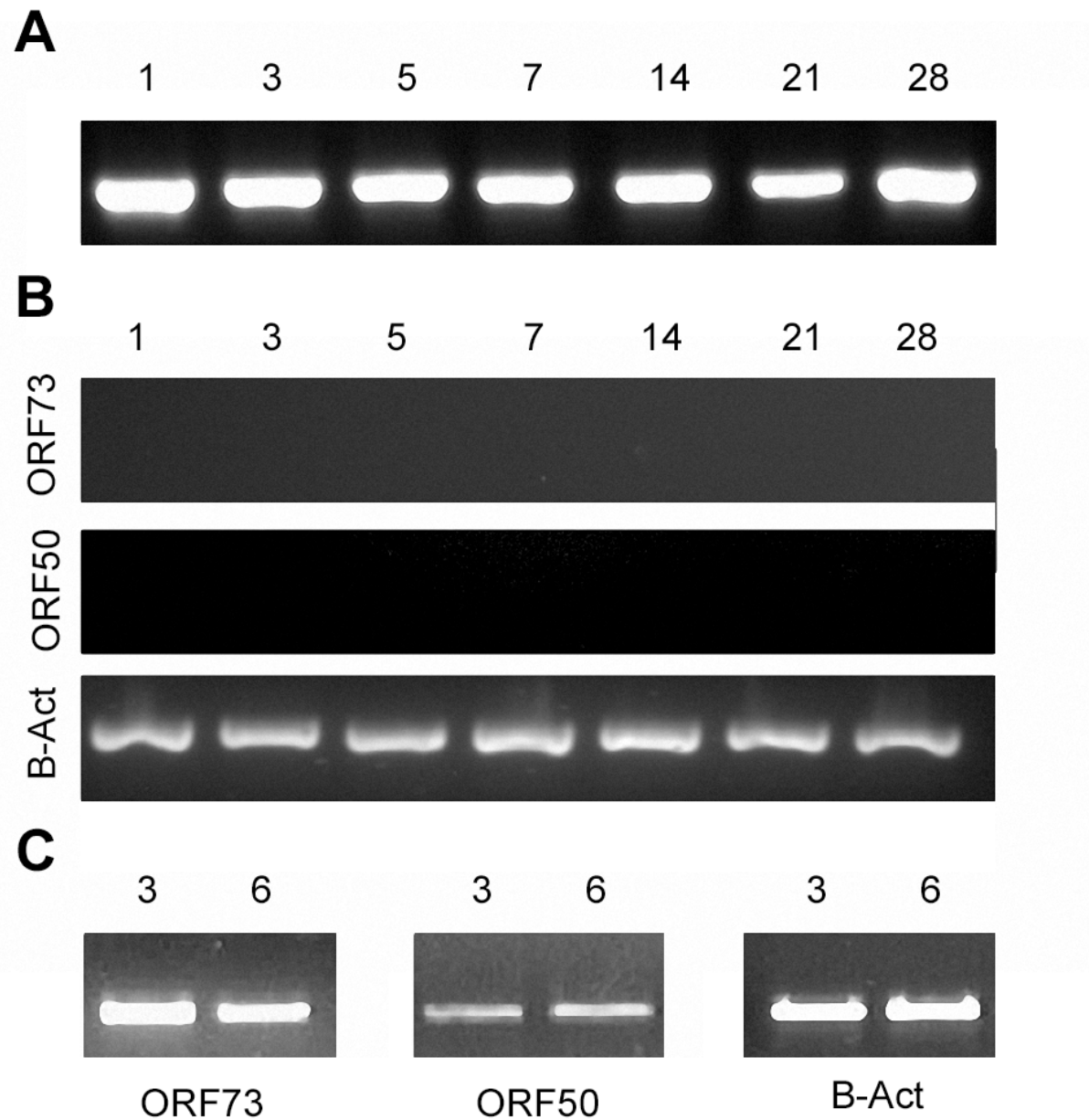


Figure 7. FM-MSCs are not permissive to HHV-8 productive infection. (A) single step qualitative PCR performed on nucleic acid extracts of FM-MSCs infected with HHV-8 at low moi (5×10^2 genomes/cell) detects a strong viral specific DNA signal at all time points tested. (B) nested RT-PCR against the latency transcript ORF73 or the lytic cycle associated transcript ORF50, on nucleic acids extracted from HHV8 infected FM-MSCs at low moi (5×10^2 genomes/cell) failed to detect HHV-8 transcriptional activity both in first and second round amplification (data not shown). (C) HHV-8 transcriptional activity is detectable by RT-PCR in high moi (5×10^3 genomes/cell) infected FM-MSCs: both ORF73 and ORF50 transcripts are amplified in single-step RT PCR until 6 days p.i

Chapter III

Turning the HIV integrase into a dis-integrase: tailoring a viral enzyme to promote the disruption of its natural target

Turning the HIV integrase into a dis-integrase: tailoring a viral enzyme to promote the disruption of its natural target

Once in a target cell, the HIV RNA genome is reverse transcribed to DNA and then undergoes integration into the host genome that is mediated by the HIV integrase (IN). IN is an enzyme with no homologue in human cells and with strong sequence specificity for HIV LTR ends. Current antiretroviral treatments do not eradicate the infection from HIV reservoirs and are hampered by the appearance of drug resistant strains. Here we present an attractive alternative which could lead to the eradication of the HIV-1 provirus from infected cells: we have conceived a novel fusion enzyme called trojIN in which IN is covalently fused via a high flexible fifteen-residue arm to the catalytic domain of FokI type II endonuclease (cdFok). In our hypothesis trojIN chimeras specifically recognize HIV-1 LTR DNA and excise it from insertion sites via its non-specific DNA cleavage domain.

Background. HIV-1 is the etiological agent of AIDS, a lifelong disease which if untreated leads to the death of the infected individual. In more than 30 years of research on the field of HIV-1, an impressive amount of progress has been achieved, and the production of new generation of anti-retroviral therapies has warranted a long life expectancy for infected people. However such therapies are not able to purge the organism from HIV-1 reservoirs, binding patients to a lifelong pharmacological treatment. The only way to free infected people from the HIV-1 plague is to clear viral reservoirs in the organism. This could theoretically be achieved in two ways: killing selectively all infected cells, or eradicating the provirus from sites of integration in their chromosomes. A way to selectively target and eliminate latently infected T cells is still unknown. In 1997 Batchu & Hinds proposed the substitution by homologue recombination of proviral DNA with an LTR-flanked therapeutic DNA, mediated by an IN-rep78 protein (1); however this fascinating hypothesis has never been translated into scientific publications. More recently Sarkar and coworkers have evolved Cre recombinase to recognize HIV-1 strain TZB0003 LTR. This modified recombinase recognizes specifically LTR and cleaves proviral DNA(2). Unfortunately, the LTR sequence of strain TZB0003 is atypical for HIV, and was chosen for these studies because of its similarity to the DNA sequence recognized by the wild type Cre recombinase. The corresponding LTR sequence from other HIV strains are much more divergent from the wild type Cre recognition sequence, and thus it is unlikely that this approach can ever be generalized to clinically relevant HIV strains. Instead to evolve a DNA binding protein to viral specific sequences, Horner and DiMaio exploited the naturally occurring HPV E2 DNA-binding activity to assemble a chimeric nuclease: they generated a fusion protein between the E2 DNA binding domain and the FokI catalytic domain (cdFok). They tested the ability of such chimeric protein on HPV infected cells, and showed a predominantly specific DNA cleavage. Actually some off-target cleavage has been detected but overall their results clearly show that a naturally occurring viral DNA binding protein could be tailored as a weapon against the virus itself (3). In our study we have selected the HIV-1 IN as the naturally occurring viral DNA binding protein. IN is composed of three domains: the N-terminal domain (aa 1-49), the catalytic core domain (aa 50-212), and the C-terminal domain (CTD; aa 213-288). The N-terminal domain has a His2-Cys2 zinc finger structure, similar to that commonly found in the DNA-binding regions of transcription factors and seems to be involved in the folding stabilization and proper multimerization of the integrase subunits (4, 5). The core domain plays a critical role in integrase enzymatic activity, includes the catalytic triad D64, D116, E152 and is involved in multimerization and in physical interaction with specific nucleotides inside LTR attachment sites

(6,7,8,9). The CTD is involved in the binding with viral and cellular DNA and has frequently been associated with unspecific DNA-binding activity (10). However several studies have shown that specific CTD residues are involved in direct interaction with specific LTR nucleotides. Over 60% of the IN aminoacidic sequence is highly conserved, and conserved residues are scattered over the entire sequence with clusters of conserved motifs, corresponding to functional sites, present in all three IN domains. It is therefore hardly possible to introduce major modifications into the primary sequence of IN without hampering its catalytic functions. However, when comparing IN sequence between different animal retroviruses the less conserved domain is the CTD (11): we predict that careful modifications altering the CTD of IN may leave its DNA specific recognition ability unaltered. Our first aim has therefore been to generate fusion proteins between native IN (and IN C-terminal mutants) and the catalytic core of FokI. We have created four different trojIN versions: one carrying the native sequence of IN, trojIN, one based on a catalytic defective variant of IN, trojIN D116A, in order to reduce possible catalytic interference by IN CCD to trojIN activity; in addition we have created two IN CTD-trojIN deletion mutants: one trojIN-delCTD in which all IN CTD coding sequence have been removed, and a trojIN version deprived of the last 18 IN aminoacid residues; the latter couple of mutants has been generated in order to remove or even attenuate IN non-specific DNA binding activity. Our final goal is the design of a novel enzyme that would act as a trojan horse to the integrated HIV genome, retaining the ability to recognize the HIV LTR as its unique DNA binding target, but once tethered to it would digest both DNA strands downstream to the binding site. We will test trojIN activity in vitro on LTR bearing plasmids, in which U3 and U5 LTR attachment sites are juxtaposed or separated by several nucleotides in order to allow the head to head accommodation of two trojIN molecules, which would result in pairing of the two cdFok on the opposing strands of interposed DNA (Fig. 1), and its subsequent cleavage.

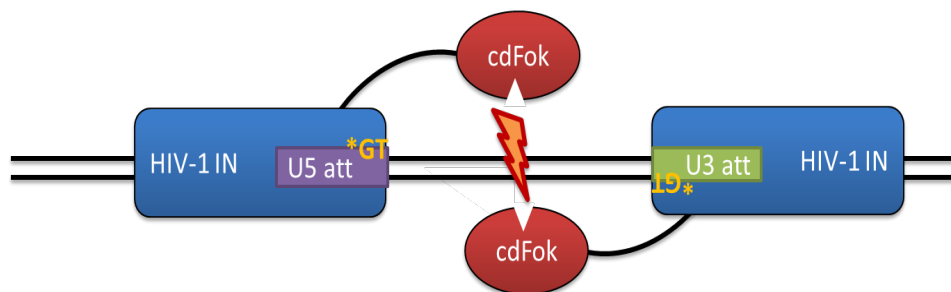


Fig. 1. Schematic representation of our experimental design. Each trojIN molecule bind to an LTR attachment site. The correct spacing between att sites will result in correct cdFok pairment which in turn leads to DNA double strand break.

Materials and Methods.

E.coli Strains and Construction of expression plasmids. Each construct was first propagated in E.coli strain TOP10. The names, sequences and main features of all oligonucleotides reported in this paper are listed in Table 1. **Native IN.** Synthetic spacer L3 has been purchased from Integrated DNA Technologies (IDT) as single strand oligonucleotides; double strand spacer has then been obtained with a fill-in reaction using the Rv L3 oligonucleotide as a primer for opposite strand synthesis. The PCR product was then cloned into the TA cloning vector pCR8 (Invitrogen) following manufacturer's guidelines, which generated the pCR8-L3 construct. cdFok cDNA was PCR amplified from the *Flavobacterium Okeanokoites* (ATCC number 40898), with oligonucleotides Fw Fok and Rv Fok. The purified PCR product has been cloned into pCR8-L3 after appropriated RE digestion of both PCR product and acceptor plasmid, in order to obtain pCR8-L3cdFok. HIV IN cDNA has been PCR amplified from a clinical isolate of HIV infected cells with oligonucleotides Fw IN and Rv IN, while L3cdFok has been amplified with oligonucleotides Fw IN-Ln and Rv Fok. Both PCR products were then put together as primers/templates in an overlapping

PCR, which resulted in the assembly of trojIN cDNA. TrojIN cDNA has then been cloned into pUC19 after RE digestion with the appropriated enzyme of both acceptor plasmid and PCR product, giving rise to pUC-TrojIN3 construct. **TrojIN mutants.** All trojIN deletion mutants (trojIN-del2, , trojIN-delCTD) have been generated following the aforementioned procedure with the exception of the oligonucleotide pairs adopted to amplify each IN mutant (see table1). Instead trojIN point mutant trojIN-D116A has been generated using the Quick Change Mutagenesis Kit (Agilent) following manufacturer's instruction (see table 1 for oligonucleotides adopted for each mutagenesis). **Gateway System.** Both native and mutated trojIN cDNA have been amplified with primers carrying attB1 and attB2 recombinant sites, using pUC-trojIN constructs as templates. Obtained PCR fragments have been inserted into the pDNR207 plasmid (Invitrogen) via BP reactions, following manufacturer's instructions. TrojIN cloned cDNA have been transferred to the gateway expression plasmids pDEST15 and pEPI-GFP, via LR recombination reaction. The att0, att20 att40 targets where obtained by single strand synthesis using as template the purchased positive strand and Rv LTR oligonucleotide as primer (table1). The PCR fragments were then gel purified and directly cloned into pCR8 as described above. The integrity of all constructs has been confirmed by DNA sequencing (PRIMM).

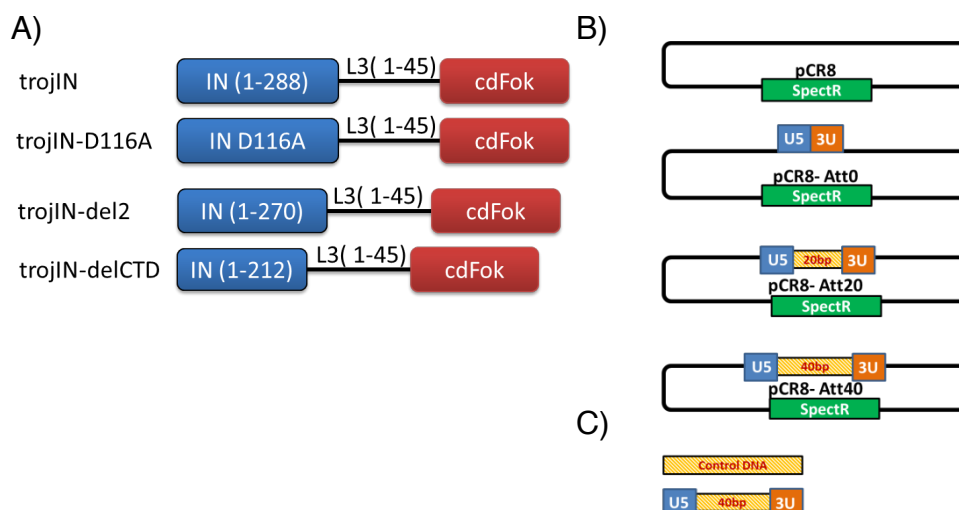


Fig. 2 Schematic representation of DNA constructs used in this work. A) TrojIN chimeras: cdFok domain (Blue) is linked by flexible L3 to the different IN subunits (Red, IN selected residues are between brackets). B) Target plasmid DNA: plasmid adopted in EMSA experiments: pCR8 is the negative control in which att0, att20, and att40 target sequences have been inserted. C) short sequences used in EMSA experiments.

Oligonucleotide	5' -> 3' Sequence	bases	RE Site(s)
Spacer L3	<u>CCGCGGGGAGGAGGAGGAAGCGGTGGAGGAGGCAGCGGCGGC</u> <u>GGTGGTAGCGCATGCTTTGGAGGTACC</u>	69	SacII, SphI, KpnI
Rv L3	<u>GGTACCTCCAAAGCATGCGCTACC</u>	24	KpnI
Rv Fok	<u>GGGGGGTACCTTAAAAGTTTATCTCGCCGTTATTA</u>	35	KpnI
Fw Fok	<u>GCCAGCATGCAAGCAACTAGTCAAAAGTGAAGTGGAGGAGA</u>	41	SphI
Fw IN-L3	<u>CAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATCCGCGG</u> <u>GGAGGAGGAGGAAGCG</u>	60	
Fw IN	<u>GGTTTAGTCGACATTTTTAGATGGAATAGATAAGGCACAAGA</u>	42	SalI
Rv IN	<u>CCGCGGATCCTCATCTGTCTACTTGCCACACAATCATC</u>	39	
Rev Lat3 pUC	<u>GGAATTGGTACCTTAAAAGTTTATCTCGCCGTTATTAAATTTCCG</u>	45	KpnI
Fw attB1 IN	<u>GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAGATAGA</u> <u>ACCATGGCGTCGACATTTTTAGATGGAATAGATAAGG</u>	80	
Rv attB2 Fok	<u>GGGGACCACTTTGTACAAGAAAGCTGGGTGGTACCAAAGTTTAT</u> <u>CTCGCCGTTATTAAATTC</u>	63	
Fw delCTD-Fok	<u>GTAGACATAATAGCAACAGACATACAACTAAAGAACCGCGGGG</u> <u>AGGAGGAGGAAGC</u>	57	
Rv IN-NoCTD	<u>CTTTAGTTTGTATGTCTGTTGCTATTATGT</u>	30	
Fw del2-Fok	<u>CAAGAAGAAAAGCAAAGATCATTAGGGATCCGCGGGGAGGAGG</u> <u>AGGAAGCGGTGGAGGAG</u>	60	
Rv IN-del2	<u>ATCCCTAATGATCTTTGCTTTTCTTCTTG</u>	29	
att0	<u>GACCGGTTGTGGAAAATCTCTAGCAGTACTGGAAGGGCTAATTTA</u> <u>CTGTGCGACGTC</u>	56	AgeI, SalI
att20	<u>GACCGGTTGTGGAAAATCTCTAGCAGTGTAGTATATGCTTGTGTT</u> <u>CCACTGGAAGGGCTAATTTACTGTGCGACGTC</u>	76	AgeI, SalI
att40	<u>GACCGGTTGTGGAAAATCTCTAGCAGTGTAGTATATGGTCATTCA</u> <u>CAAGACTCTAGTCTTGTGTTCCACTGGAAGGGCTAATTTACTGTC</u> <u>GACGTC</u>	96	AgeI, SalI
RV LTR	<u>GACGTCGACAGTAAATTAGCCCTTCCAGT</u>	29	SalI
Fw IN116A	<u>GCCAGTAAAAACAATACATACAGCAAATGGCAGCAATTCACCA</u> <u>G</u>	45	
Rv IN116A	<u>CTGGTGAAATTGCTGCCATTTGCTGTATGTATTGTTTTACTGGC</u>	45	

Tab. 1 DNA oligonucleotides used in this work

Prokaryotic protein expression and purification. GST-tagged or tag-less trojIN molecules were expressed in E.coli BL21 (DE3) (Invitrogen) strain under standard conditions: about 24h before induction a pre-culture of LB medium with trojIN expressing BL21 cells was set up and grown at 37°C with constant swirling for 16-18h. After incubation an aliquot of growing culture was inoculated in fresh media at a ratio (v/v) of 1:50, and grown in the same conditions, until culture reached an OD of about 0,6-1,0. Then protein expression was induced by adding to the culture medium IPTG at a final concentration of 0,5mM. Induced culture were harvested 6h post induction by centrifugation at 200xg at 4°C for 30 minutes. Collected cultures have been stored at -20°C until needed. GST tagged trojIN were purified by Glutathione-Sepharose Beads (Amersham) following producer's suggestions. Tag-less trojIN proteins were purified from inclusion bodies essentially as reported previously (12): briefly harvested cells were resuspended in

lysis buffer, subjected to mechanical and enzymatic lysis; then inclusion bodies were isolated by centrifugation at 12000xg for 30 minutes, washed twice with a 4M Urea solution and then solubilized with a SDS based denaturing solution for 16h. The whole protein mixture has then been subjected to 8% SDS-Page. The correctly sized protein was visualized in the acrylamide gel after CuCl_2 negative staining, the corresponding gel band was excised and subjected to electroelution in SDS-running buffer. Purified trojIN were refolded upon gradual removal of denaturing agents by step dialysis. The yield and integrity of purified proteins were assayed by SDS-PAGE followed by 1% Brilliant Blue Coomassie Staining or Western Blotting with mouse anti-IN monoclonal antibody (Santa Cruz). Purified proteins were stored in 50% glycerol buffer at -20°C .

EMSA. Purified trojIN variants activity and correct folding has been tested in DNA binding assay: 3,5pmol of trojIN were incubated with either 0.05pmol of plasmid DNA or 1,5pmol of 100bp long dsDNA for 1h at 37°C or 24h in different Tris-based buffers. Plasmids/trojIN macromolecular complexes were resolved by 0,8% agarose gel electrophoresis run 1h at 100V in 0,5x TAE buffer (Tris 20mM, Acetic Acid 10mM, EDTA- Na_2 0,5mM). Short DNA/trojIN complex were resolved in 8% acrylamide gels run at 100V for 1h in TBE buffer (Tris 45mM, Boric Acid 88mM, EDTA- Na_2 0,2mM).

Results.

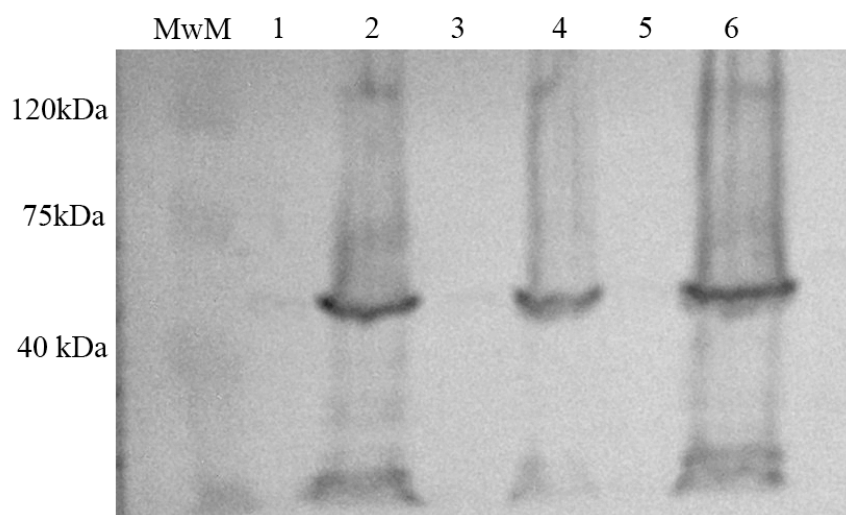


Fig. 3 TrojIN solubility evaluation. Western Blot Analysis of differentially induced trojIN expression, in order to obtain native soluble chimeric protein. MwM, Molecular Weight marker; Lanes 1, 2 trojIN expression was induced for 5h at 37°C ; Lanes 3,4 trojIN induction have been carried out at 4°C for 72h; 5, 6 trojIN expression induced for 16h at 20°C . Lanes 1, 3, 5 Native soluble fractions; Lanes 2, 4, 6 total lysate

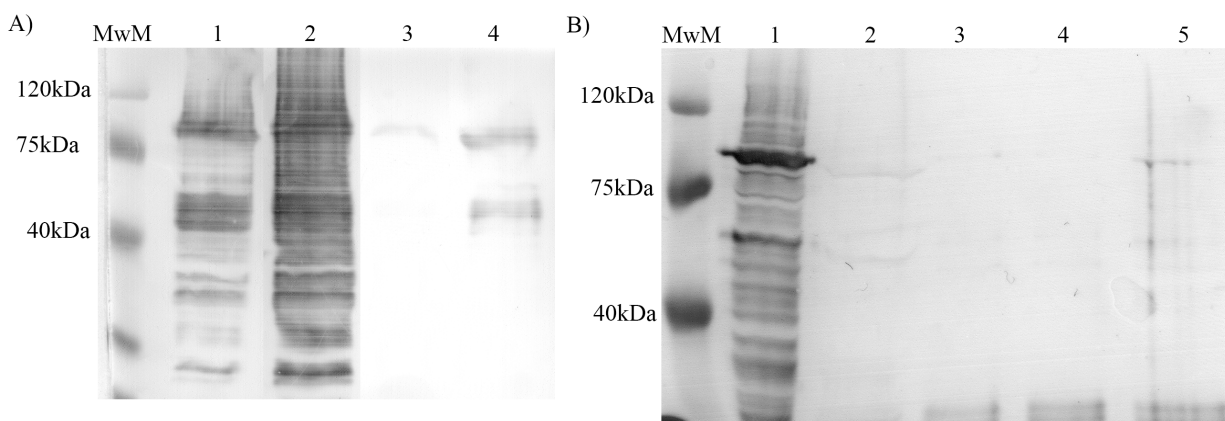


Figure 4. GST-trojIN expression and purification. A) GST-trojIN chimera expressed in BL21 cells was partially extracted from cells with a native lysis buffer and trojIN expression was detected with anti IN antibody: Lane 1 Un-induced BL21, total lysate; BL21 total lysate after 5h induction at 37°C; Lane 3 soluble fraction after non denaturing cell lysis; Lane 4 soluble fraction after 4M urea wash of insoluble material. B) Soluble GST-trojIN protein have been subjected to purification on Glutathione-Sepharose Beads columns and obtained fractions were analyzed with Western Blot developed with anti IN antibody: Lane 1 Lysate soluble fraction; Lane 2 Column flow trough (unbounded fraction); Lane 3 and 4 low Glutathione concentration eluted fraction. Lane 5 Eluted Fraction.

WB and purification. We have overexpressed several trojIN variants (summarized in fig. 2A) in E.coli BL21 strain, a suitable strain for protein overexpression, in which the main prokaryotic proteases has been genetically shut down. The expression was induced by adding IPTG and by incubating culture at different times and temperatures. We first expressed trojIN tag-less, in order to avoid fusion partner interference to trojIN activity. We have assayed different inducing conditions in order to improve trojIN solubility but first attempts to obtain tag-less soluble trojIN molecules in detectable amounts by Coomassie staining or WB analysis failed (Fig. 3). Next we assayed the solubility of GST-tagged trojIN molecules, a fusion partner largely used to enhance solubility and purification yield of overexpressed proteins (13). As shown in figure 4 GST fusions led to an increased solubilization of trojIN: chimera soluble fractions were more than barely detectable (figure 4A, LANE SN1); furthermore an abundant amount of fusion protein became soluble after treatment with a low concentration Urea buffer (see fig. 4A LANE SN2) suggesting that GST fusion allowed formation of almost soluble trojIN. We have proceeded further purifying GST-tagged trojIN as described in material and methods section. WB analysis of purification steps showed only a faint band in the “Eluate3” lane, at the expected molecular weight (figure 4B, lane E3); moreover all three elution steps showed marked degradation products. Furthermore a visible signal was detectable in the fraction corresponding to unbound fraction. These results were suggestive of a not properly corrected folding of GST-trojIN in the case of unbound fraction, and of an overall instability of the fusion products, highlighted by a high level of degradation, even in the presence of protease inhibitors. Parallel purification with another GST-fusion protein (HCMV UL54) has strengthened our conclusion: in fact induction and purification in the same conditions led to an appreciable amount of purified control protein (data not shown). In order to obtain high level of stable purified protein we next have extracted tag-less trojIN from inclusion bodies: correct sized trojIN were then isolated from denaturing PAGE and refolded under native conditions. As shown in figure 3 the above purification strategy has allowed us to obtain high level, high pure tag-less trojIN molecules.

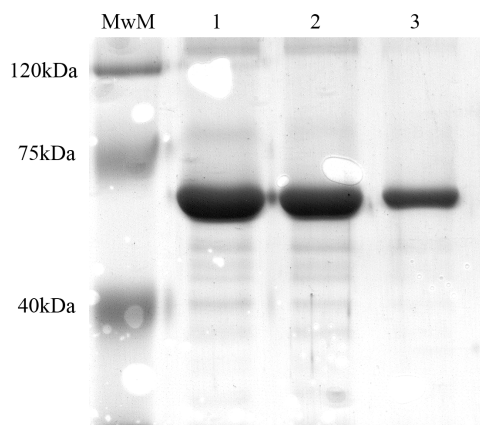


Fig. 5 Tag-less trojIN purification from inclusion bodies. BL21 expressing Tag-less trojIN were lysed as described in material and methods section. Purification steps were resolved on 8% SDS-PAGE and detected with Coomassie Brilliant Blu staining: Lane 1 Inclusion bodies total extract; Lane 2 trojIN inclusion bodies after UREA wash and 1% SDS denaturation; Lane 3

EMSA (1) plasmid. Since it has been previously shown that HIV-1 IN is able to bind LTR-carrying plasmids (14), we have assayed the ability of each trojIN version to bind linearized plasmid DNA carrying or not HIV-1 LTR attachment sites (see target section fig. 2), by EMSA on agarose gels. Besides incubation with LTR bearing DNA and functional trojIN molecules could result in some kind of cleavage of target DNA molecule. First we assayed the ability of all trojIN variants to bind to specific (with LTR) or non-specific plasmid (fig 4A-B): while native trojIN, the point mutant D116A and the negative control (protein storage buffer) seemed not to influence DNA electrophoretic mobility, trojIN deletion variants (trojIN-del2, trojIN-delCTD) appeared to bind both LTR carrying or negative control plasmids, slowing down their migration speed. Anyway under no circumstances trojIN-DNA incubation resulted in detectable cleavage of a DNA fragment. Next we have tried to optimize buffer composition with the trojIN version del2 in order to devise optimal conditions for LTR-specific binding: as shown in figure 4C, trojIN-del2 confers mobility shifts to LTR bearing plasmids but not to control DNA in buffer NEB2. We then have performed EMSA assays varying the incubation times up to 24h), but failed to detect DNA digestion products (data not shown). All above described EMSA assays have been carried out in the presence of high zinc ion concentration (50 μ M); it is well known that free zinc is required for proper IN dimerization/multimerization and to enhance Mg²⁺ dependent IN activity (15,16,5). In order to evaluate if the observed DNA binding was correlated to IN multimerization, we have performed EMSA assays with varying Zn²⁺ concentrations: as illustrated in figure 4D trojIN retained its ability to bind DNA in standard Zn²⁺ concentration as seen in previous experiments. In contrast EMSA performed in reaction buffer with a 5 fold reduction in zinc concentration or without Zn²⁺ ions, showed abolished trojIN induced mobility shift.

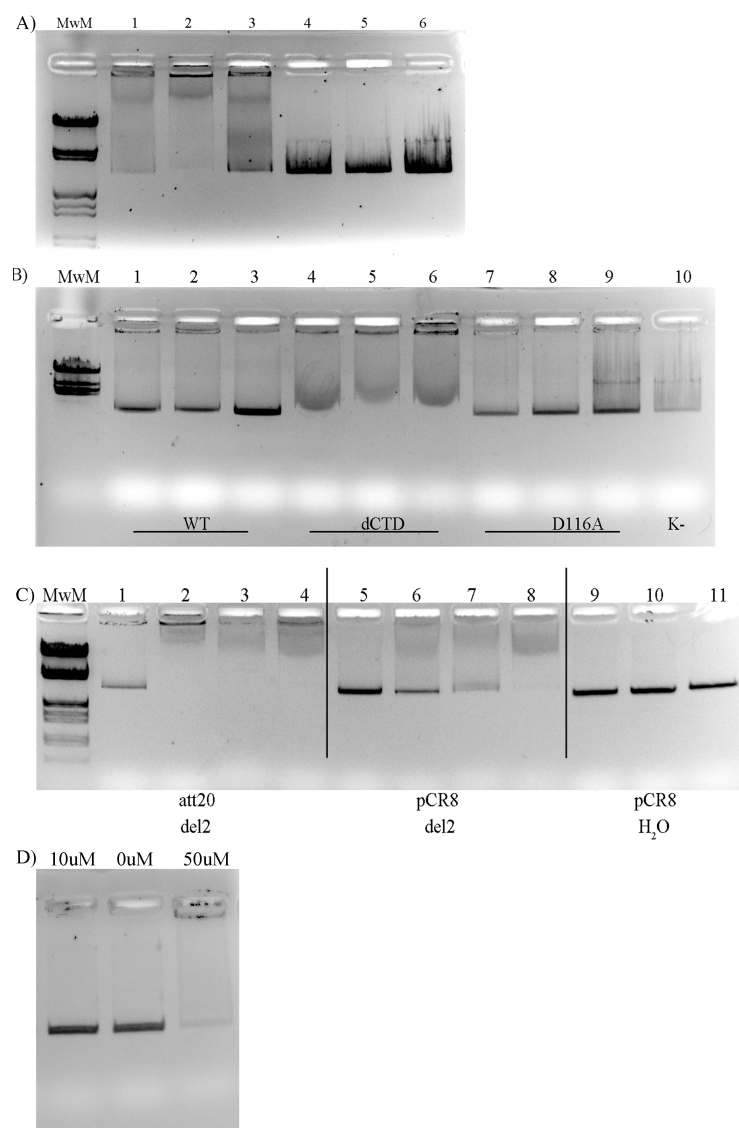


Fig 6. Agarose EMSA with trojIN variants, trojIN specific DNA binding has been evaluated under different saline conditions.

A) Lanes 1, 4 pCR8 plasmid; Lanes 2, 5 att0 plasmid; Lanes 3, 6 att40 plasmid. trojIN-del2 (lanes 1-3) incubated with specific or aspecific DNA target in buffer 2 (TrisHCl 10mM pH 7.9, NaCl 50mM, MgCl₂ 10mM, DTT 1mM) was run on 0,8% agarose gel. In lanes 4-6 samples are Dialysis buffer incubated with DNA targets as negative controls.

B) trojIN WT (lanes 1-3), -delCTD (lanes 4-5) and -D116A (lanes 7-9) trojIN variants or dialysis buffer (lane 10) incubated with pCR8 (Lanes 1, 4, 7, 8) att0 (Lanes 2, 5, 8) or att40 (Lanes 3, 6, 9) were run on 0,8% agarose gel.

C) trojIN-del2 variant was incubated with att20 or control plasmids in different saline buffer conditions. Lanes 1, 5 Buffer 1 (Bis-tris-propane HCl 10mM pH 7.0, MgCl₂ 10mM, DTT 1mM); Lanes 2, 6 Buffer 2; Lanes 3, 7 Buffer 3 (TrisHCl 50mM pH7.9, NaCl 100mM, MgCl₂ 10mM, DTT1mM); Lanes 4, 8 Buffer 4 (TrisAcetate 33mM pH 7.9, K-Acetate 66mM, Mg-Acetate, DTT 1mM). As control Dialysis buffer (H₂O) was incubated in teh same conditions (Lane 9 Buffer 1, 10 Buffer2, 11 buffer3)

D) EMSA after trojIN-del2 interaction with att20 plasmid in buffer 2 with different free zinc concentrations

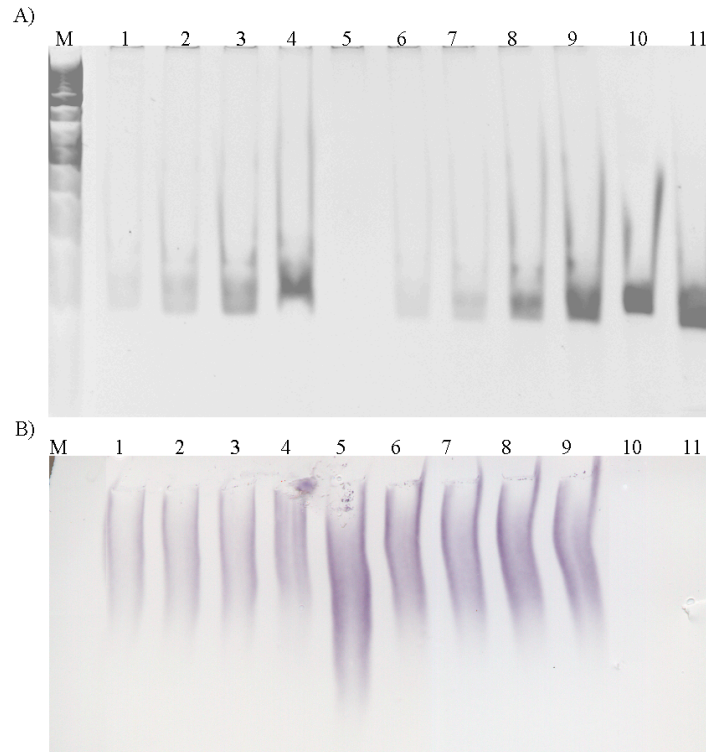


Fig. 7 EMSA after trojIN-del2 incubation with att40 or control Short DNAs.

A) Etidium bromide staining of TBE-PAGE 8%. TrojIN-del2 (Lanes 1-9) incubated with buffer2 (lane 5) or with different amount of att40 DNA(Lanes 1-4) or control DNA (Lanes 6-9). Lanes 1, 6 0,15 pmol of DNA; Lanes 2, 7 0,38pmol of DNA; Lanes 3,8 0,75pmol of DNA; Lanes 4, 9 1,5pmol of DNA. Lane 10 att40, Lane 11 Control DNA.

B) Migrated DNA/protein complexes were transferred on Nitrocellulose subsequently subjected to Western Blot analysis with anti IN antibody

EMSA (2) short DNA. EMSA has also been assayed with low molecular DNA target carrying LTR attachment sites, called att40 from now on, or with a same length control DNA; we have adopted the same optimized conditions determined above. Figure 5B shows clearly that IN binds both to specific or non-specific DNA: electrophoretic mobility of trojIN-del2 protein, incubated in reaction buffer in the absence of DNA (lane5) is higher if compared with DNA associated trojIN-del2 complex (lanes 1-4 and 6-9). It appears unequivocal that DNA-trojIN interaction resulted in a high molecular weight complex formation. Furthermore as observable in fig 5a-lane 4, trojIN-del2 protein reduced DNA electrophoretic mobility in a LTR specific fashion since a shift in control-DNA electrophoretic mobility was not detected (compare lane 4, att40 with lane9, control DNA). However DNA/trojIN complex formation seem less efficient on short DNA substrates compared with plasmid DNA: only the addition of high amount of DNA substrate (1,5pmol, fig.5A lane 4) allowed trojIN-del2 to impose mobility shift as observable comparing lane 4 with lanes 1 to 3 (Fig. 5a). This results suggests that on short DNA an high trojIN:DNA ratio could inhibit complex formation. Anyway at any rate trojIN incubation with short DNAs resulted in DNA digestion.

Discussion and future perspectives.

HIV establishes lifelong latency within infected memory T cells, and current therapies cannot eradicate the infection, even if they successfully inhibit viral replication. The only definitive cure for the HIV infection lies in the complete elimination of latently infected cells or of the integrated provirus they carry. Recent works have attempted to remove the T cell compartment by non-specifically activating HIV within memory T cells, hoping that activated viral replication would result in the disruption of latently infected cells (17,18). But such promising approaches have to date proven unsuccessful. Furthermore it is hard to predict the effects of a widespread T-cell disruption consequent to targeting HIV latently infected T cells for destruction, on the ability of the infected individual to reconstitute an effective immune response. Alternatively an attractive approach would be to selectively target the integrated provirus within these cells. Our work belongs to this novel field of research, a preliminary and rather incomplete attempt to generate a molecular tool for the excision of the HIV provirus from insertion sites: we have generated four different trojIN candidate enzymes, namely trojIN, trojIN-D116A, -del2 and -delCTD. We have tested the ability of these novel chimeras to bind LTR bearing DNA. As a result we have identified two trojIN variants -del2 and -delCTD, which bind electively (-del2) or with more efficiency (-delCTD) the LTR bearing pCR8-att20 plasmid. In addition trojIN-del2 apparent preference for LTR bearing DNA has been confirmed by EMSA experiment performed with the 100bp-long LTR carrying att40 DNA (Fig5a,b): trojIN-del2/att40 interaction with a 1:2 molar ratio resulted in a lowered DNA electrophoretic mobility, which is not detectable in the case of trojIN-del2/control DNA interaction. These results are consistent with our working hypothesis, namely that suitable modifications of the CTD sequence may reduce IN binding to non-specific DNA. However the reason for full-length trojIN inability to bind DNA in our hands remains elusive, since IN protein is able to bind plasmid DNA *in vitro*(14). One possible explanation could reside in the presence of cdFok which could interfere with the correct folding of full length IN. Anyway the main result is at any rate the lack of a detectable cdFok activity, and it is possible that DNA cleavage products are not detected in our experimental conditions due to very low efficiency of the reaction. Another explanation could be an inappropriate trojIN complex formation: the rationale behind target design was to mimic naturally occurring 2LTR circles, an hypothesized viral DNA reservoir in infected cells (19), which has been demonstrated to be recognized in *in vitro* experiments by recombinant IN. We have introduced spacer nucleotides between LTR att sites in order to allow proper distancing between LTR- bounded trojIN molecules: it is well known that cdFok cleaves DNA upon dimerization with a partner cdFok domain (20) and a correct trojIN distancing would allow proper cdFok dimerization which in turn would result in DNA cleavage (Fig. 8A). On the contrary if spacing between trojIN molecules would not be correct, namely LTR att sites are too far or contiguous cdFok subunit couldn't interact suitably, resulting in lack of cleavage activity (Fig. 8B). Moreover trojIN dimerization could occur between IN surfaces (Fig. 8C/D), resulting in a constrained trojIN dimer, potentially unable to be catalytically active.

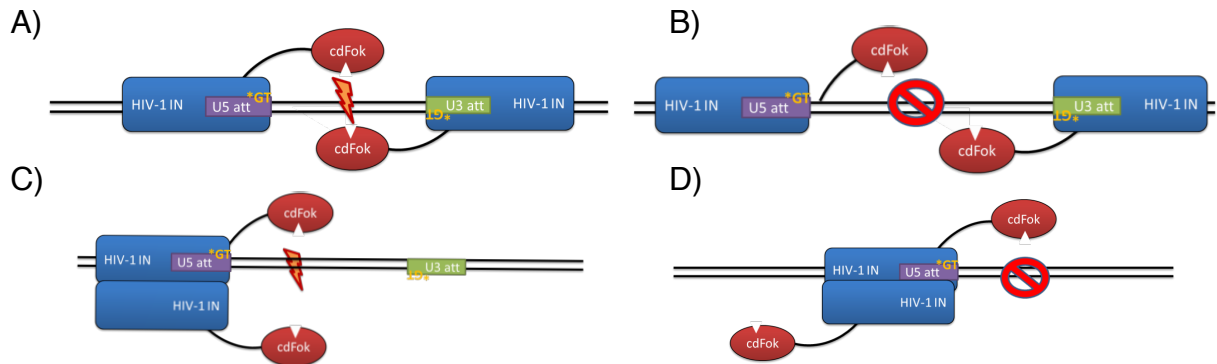


Fig. 8 Schematic representation of hypnotized trojIN molecular rearrangement on target DNA.

A) Correctly spaced trojIN pair allow proper cdFok dimerization and subsequent DNA cleavage, which couldn't take place if LTR sequences are too far from each other (B). C) Also dimerization through IN interface could lead the proper pairing of two cdFok domains, but could also result in physical separation of the two cdFok domains (D).

The actual trojIN dimerization, if occurring, would not be easily predictable, but we can try to minimize un-functional trojIN dimer formation: it has previously characterized by others (21) a IN single-aminoacid mutant with impaired IN omo-dimerization; adopting a trojIN version carrying such mutation we could reduce trojIN un-functional dimerization. However at the moment we are working into two main directions in order to obtain more functional chimeras: one is the creation of trojIN variants with a lengthened flexible arm connecting IN and cdFok, which could overcome eventually occurring structural constrain between the two trojIN domains. Another possibility to enhance trojIN specific activity is to set up an *in vivo* evolution system: we are now working on an adapted *in vivo* evolution system performed in prokaryotic cells and inspired by the work of Barbas III's group (22): briefly we have inserted our targets DNA (att0, att20, and att40) into a toxic gene (ccdB) expressing plasmid, and with them transformed BW25141 strain, an engineered E.coli strain capable of a rapid degradation of linear DNA (23). The induction of ccdB gene expression leads cells to death, warranting a stringent selection system. Subsequently we have created a cDNA library with randomly mutated trojIN, which have been subjected to recombination via DNA shuffling and subsequent cloned into an expression plasmid. TrojIN plasmid have been used to transform the BW25141 strain of E. coli, carrying the LTR-ccdB plasmid, and the expression of both toxic gene and trojIN have been induced: only clones transformed with trojIN having some degree of activity have grown, representing the first generation of selection. DNA plasmid extracted from selected colonies have then be used as templates to amplify trojIN with error-Prone PCR, creating a new library of mutated chimeras. Finally we have repeated recombination and *in vivo* selection obtaining trojIN second generation. Preliminary results are promising: in the first generation we have obtained about ten colonies, while in the second about two times suggesting an increased trojIN efficiency. Furthermore the preference for att20 target plasmid observed in EMSA experiments (Fig.4) seem to occur even in the selection system, supporting our *in vitro* results. Functional trojIN variants will be transferred in eukaryotic reporter system: we plan to create a fluorescent reporter cell line transduced with a GFP expressing lentiviral vector. TrojIN activity will be measured after transfection of the reporter cell line with an eukaryotic trojIN expression vector: cytofluorimetric analysis over time of the number of GFP expressing cells will be the measure of trojIN specific activity. Furthermore cytofluorimeter will allow us to determine the cell cycle phase of analyzed samples: for example a cell population with reduced amount of population and in a SubG0 phase will be suggestive of apoptotic events, probably due to trojIN cytotoxic activity or trojIN off-target activity. The resulting positively selected trojIN variants will be *bona fide* candidate for moving toward pre-clinical research phase.

References

31. Batchu & Hinds ' 97, Med Hypotheses 69: 967
32. Sarkar et al' 07, Science 316: 1912
33. Horner&DiMaio, J. Virol, 81:6254.
34. Bushman et al' 93, Proc. Natl. Acad. Sci. USA, 90:3428
35. Zheng et al' 96, Proc. Natl. Acad. Sci. USA, 93: 13659
36. Dolan et al' 09, J. Mol. Biol. 385: 568
37. Esposito and Craigie ' 98, EMBO J, 17: 5832
38. Gerton and Brown ' 97, J Biol. Chem. 272: 25809
39. Jenkins et al' 97, EMBO J, 16: 6849
40. Lutzke et al' 97, Nucleic Acid Res., 22: 4125
41. Cathomen and Joung ' 08, Mol. Ther. 16: 1200
42. Falconer et al' 97, Biotechnol. Bioeng. 57: 381
43. Ghosh et al' 95, Anal. Biochem. 225: 37
44. Delelis et al' 07, PLoS One 2: e608
45. Deprez et al '00, Biochem. 39: 9275
46. Lee et al' 97, Biochem 36: 173
47. Lehrman et al' 05, Lancet 366:549
48. Siciliano et al' 07, J. Infect Dis 195:833
49. Panganiban et al' 84, Cell 36: 673
50. Kim & Chandrasegaran ' 94, Proc. Natl. Acad. Sci. USA, 91: 883
51. Berthoux et al' 07, Virology 364: 227
52. Guo et al '10, J. Mol. Biol. 400: 96
53. Kuzminov and Stahl' 97, J. Bacteriol. 179: 880